1994

Use of Uncertainty Methodology in Identification and Classification of Soils Based Upon CPT.

Zhongjie Zhang
Louisiana State University and Agricultural & Mechanical College

Order Number 9425236

Use of uncertainty methodology in identification and classification of soils based upon CPT

Zhang, Zhongjie, Ph.D.

The Louisiana State University and Agricultural and Mechanical Col., 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

USE OF UNCERTAINTY METHODOLOGY IN IDENTIFICATION AND CLASSIFICATION OF SOILS BASED UPON CPT

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Civil Engineering

by
Zhongjie Zhang
B.S., Xian's Institute of Highways, Shannxi, 1982
M.S., Tongji University, Shanghai, 1984
May 1994

ACKNOWLEDGMENTS

The research described in this dissertation was financially supported by the Louisiana Transportation Research Center (LTRC), under Grant No. 88-1 GT (State Project No. 746-13-36, Louisiana HPR No. 0010(12)). This support is gratefully acknowledged.

I wish to express my thanks to Dr. Jean-Lou Chameau and Dr. Carlos Santamarina for their spending valuable time to carefully read my research proposal and for their important and critical comments on this research.

This sincere appreciation is also extended to the people who provide in-situ cone testing data and boring log data. They are Dr. J.-L. Briaud from Texas A&M University, Dr. M. W. O'Neill from University of Houston, Dr. J. R. Martin, II and Dr. W. Clough from Virginia Polytechnic Institute and State University, Dr. J. K. Mitchell from University of California, Berkeley, and Dr. An-Bin Huang from Clarkson University, Mr. P. Griffin and Mr. J. Oglesby from LTRC, and Dr. Jim Laice from Southern Earth Science, Inc. Their contribution has made it possible to finish this research. Specifically, Mr. P. Griffin and Mr. J. Oglesby are deeply appreciated for supporting and helping the research work.

Fundamentally, I am deeply indebted to Dr. M. T. Tumay, my major professor and advisor, for his invaluable inspiration, constant encouragement, understanding, and important guidance during the whole process of this research. Also, his friendship to me has alleviated the challenging student career I spent at Louisiana State University (LSU).

My appreciation is extended to other professors of my advisory committee: Dr. J. Chen, Dr. Y. Acar, Dr. G. Z. Voyiadjis, and Dr. J. Suhayda for their critical comments and suggestion on this research. I also would like to express my admiration and gratitude for their guidance and encouragement when I took graduate courses from them. Moreover, Dr. Voyiadjis and Dr. Acar are thanked for providing computer facilities to let me finalize my dissertation.

A special thank should go to Mr. Bill Tierney for his assistance and help to me during my study and research at LSU. His friendship will be cherished through my life.

Finally, a very special thank should be given to my whole family, especially to my lovely wife Shaoqing Guo, dear daughter Xinyuan Zhang, and my respected parents. Their love, supports, and encouragement have continuously been one of the resources to let me finish my study for PhD at LSU. This dissertation is also a welcome gift to my dear newly-born son, David Zhang.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

A	Classic set
a	Net area ratio of a cone
B, B_α	Fuzzy subset
B_q	Equal to $(u - u_0)/(q_t - \sigma_{v0})$
CFSC	CPT Fuzzy Soil Classification
CL	Lean clay
CH	Fat clay
CPT	Cone Penetrometer Test
CPTCC	CPT Classification Chart
CPTU	Piezo-Cone Penetrometer Test
CST_j	jth Crisp Soil Type
c	Clayey soils
d_0	Middle point of a moving window
$F_1, ..., F_7$	Cumulative distribution function for different soil types
F_R	Equal to $f_s/(q_t - \sigma_{v0})$, 100%
FR	Friction Ratio
FR_1	Corrected friction ratio
$F_m(U)$	Marginal distribution function of U with respect to F(U, SI)
F(U, SI)	Two dimensional distribution function of U and SI
$f_c(u)$	Density function for HPC
$f_m(u)$	Density function for HPM
$f_s(u)$	Density function for HPS
f_s	Cone friction resistance
f_{s1}	Corrected cone friction resistance
HPC	High Probability Clayey Soils
HPM	High Probability Mixed Soils
HPS	High Probability Sandy Soils
ICC	Intraclass Correlation Coefficient
GP	Poorly graded gravel
ML	Silt
M	Number of soil types
m	Equal to $(\mu_1 - \mu_2)^2/(\sigma_1^2 + \sigma_2^2)$
m	Mixed soils
N	Number of colors
n	Exponent value

n	$n_1 + n_2$, the size of the combined sample consisting of Ω_1 and Ω_2
n_1, n_2	Sizes of samples Ω_1 and Ω_2, separately
OCR	Overconsolidation Ratio
PCPT	Piezo-Cone Penetrometer Test
$p_{i,j}$	Probabilities with which each classification region receives each type of soils
Q_t	Equal to $(q_t - \sigma_{v0})/\sigma_{v0}'$
$q_{i,j}$	Probabilities with which each type of soil falls in each classification region
q_c	Cone tip resistance
q_{c1}	Corrected cone tip resistance
q_{si}	Marginal distribution function of SI with respect to F(U, SI)
q_t	Corrected total cone resistance
R_1, ..., R_7	Soil classification regions corresponding to different soil types
R_f	Friction Ratio
RSTDFR	Relative STD of Friction Ratio
RSTDTR	Relative STD of Tip Resistance
SC	Clayey sand
SI	Soil type index
SM	Silty sand
SP	Poorly graded sand
STD	Standard Deviation
s	Sandy soils
$U_{i,j}$	Boundary values among soil classification regions, i, j = 1, ..., 7
USCS	Unified Soil Classification System
U	Soil classification index
U, u	Soil in-situ behavior unit
u, v	Coordinates in W plane
u, u_0	Pore pressure between the cone tip and the friction sleeve
V	In-situ state index
W	W complex plane
W_d	Width of a moving window
w	Complex variable in W plane
x, y	Coordinates in Z plane
Z	Z complex plane
z	Complex variable in Z plane
α	Angle between two curves

α	Significance level
α, α_i	Level value for a fuzzy subset
μ	Mean of the combined sample consisting of Ω_1 and Ω_2
μ_1, μ_2	Means of samples Ω_1 and Ω_2, separately
$\mu_A(x)$	Membership function of a classic set A
$\mu_B(x)$	Fuzzy membership function of a fuzzy subset B
$\mu_c(u)$	Fuzzy membership function for HPC
$\mu_m(u)$	Fuzzy membership function for HPM
$\mu_s(u)$	Fuzzy membership function for HPS
ρI, ρI_{peak}	Intraclass correlation coefficient
σ_1^2, σ_2^2	Variances of samples Ω_1 and Ω_2, separately
σ_v', σ_{v0}'	Effective vertical stress
σ_{v0}	Total vertical stress
Y_b^2	Between class variance, the variance of the combined sample consisting of Ω_1 and Ω_2
Y_w^2	Pooled combined variance
Ω_1, Ω_2	Data samples on each side of the window center d_0, separately
$\in$	Means "belong to"
$\cap$	Means "coexist"
$\notin$	Means "not belong to"
$\forall$	Means "all"

ABSTRACT

The current Cone Penetration Test (**CPT**) soil engineering classifications have two kinds of uncertainties: randomness and fuzziness. Research indicates that possible solutions to these uncertainties can only be worked out through modeling them. Therefore, a systematic investigation is performed and some preparative tasks are done in advance. First, an efficient soil classification index, U, is defined and several **CPT** soil classification charts are simplified accordingly. Second, a moving window approach based upon an Intraclass Correlation Coefficient (**ICC**) to determine normal soil behavior units is adopted so that a correlation between soil types and soil behavior units can be established. Based upon these approaches, a preliminary data reduction is performed on the raw **CPT** data from eight sites, and the characteristics and distributions of the soil behavior units are determined and discussed for seven soil types.

Two statistical criteria, Region Estimation and Point Estimation, based upon distributions of soil behavior units are then developed to predict soil type using **CPT** data. Also, a fuzzy subset approach is suggested to handle the fuzziness and randomness. In this **CPT** fuzzy soil engineering classification, a new naming system is used. The randomness of **CPT** soil engineering classification is put into the conceptual framework of three new soil types. The fuzziness is then described by fuzzy membership functions. These functions are derived from the modification of the density functions of corresponding compositional soil groups.

Finally, a new package of **CPT** soil engineering classification is suggested. It consists of following procedures:

1). Transform a **CPT** sounding profile of parameters (tip resistance, q_c, and friction ratio, FR) by conformal mapping to a corresponding profile of soil classification index, U;

2). Layer the U profile by **ICC** moving window method and calculate the mean of U values for each layer to determine the soil behavior unit;

3). Predict the soil type cf each layer by matching the soil behavior unit of that layer with the classificaticn criteria suggested in this study.

Several sets of **CPT** soil engineering classification criteria are recommended in this dissertation. They are the indicators of an evolution process from the purely empirical to the purely theoretical.

INTRODUCTION

The identification and classification of in-situ soil types have become one of the primary applications of Cone Penetrometer Test (CPT). Since 1965, when first soil classification chart based on the results of Begemann mechanical friction cone came to the literature, this topic has been kept as one of the concentrated areas for research. Some reasons are

- This technology has been considered as ideally suited for use in site investigations and profiling;
- The knowledge of soil types, such as sand, clay, etc, helps the interpretation of cone data on the corresponding soil engineering properties;
- No soil sample is available from CPT for visual and laboratory inspection;
- The ample information and knowledge on soil engineering properties and behaviors have been accumulated in terms of existing soil classifications of compositional type, such as the Unified Soil Classification System (USCS).

However, most importantly, the prediction of soil types by current soil classification charts of this technology sometimes do not match the real situation of in-situ soils well, although in recent years these charts have been adapted and improved substantially from an expanded database. It is believed that all the charts to date can not be expected to provide accurate prediction of soil types from a compositional classification, but only a guide to a soil classification of behavior type (**Douglas & Olsen, 1981; Campanella**

& Robertson, 1988). Then, what is a soil classification of behavior type? What is its nature? What are its differences and relations with conventional soil classifications of compositional type? The answers to these questions will help us to get the insight to **CPT** soil engineering classification and to correctly interpret and use the soil classification data from cone penetration technology. Therefore, it is imperative to perform a systematic investigation on the problems behind those questions. Only based upon this kind of effort, could a feasible solution for the mis-prediction problem be found out. This dissertation has been motivated in such an attempt to explore the insight to the general soil engineering classification problem. The outcomes of this systematic research are expected to provide a sound foundation for a possible **CPT** soil engineering classification to suggest.

The results of this investigation will be arranged and presented from both theoretical and practical points of views. Such an arrangement will also reflect the technical route of this research to follow. There are three parts in this dissertation. The first part (Chapter 1 and 2) will discuss the problems of soil classification indices and soil stratigraphy. These two issues are closely related in that they are the two basic aspects to distinguish in-situ soils based upon same sets of cone testing data. They are important since their solutions will consist of the fundamental steps in a possible **CPT** soil engineering classification. An efficient soil classification index, instead of a current two-dimensional-chart format of a **CPT** soil classification, not only is easy to be implemented in a computer system but also will well preserve the advantage of continuously describing in-situ soil conditions. Such an index can also greatly reduce the difficulty and complexity of a possible theoretical analysis. Similarly, a proper procedure for in-situ soil stratigraphy will pave the way to find out the representative correlations between the composition and behavior of soils subject to cone penetration. Therefore, they are discussed first and together. Accordingly, several existing **CPT** and **PCPT** (Piezo-Cone Penetrometer Test) soil classification charts will be simplified. The results then will be checked by some in-situ testing data to show their validation. Due to the in-situ testing data available, most efforts in this research will be focused on **CPT**.

The second part (Chapter 3 and 4) of this dissertation will explore the reasons why **CPT** soil engineering classifications have an uncertainty of randomness. This uncertainty shows itself as the overlaps among different types of soils in current soil classification charts. Many causes will be found to account for this phenomenon. As pointed out before, this uncertainty can result in mis-prediction on soil types. In order to conquer such a problem, the possibility for a proper modelling of it will be studied. Some probable approaches to model the uncertainty will be fully developed and some in-situ cone testing data will be used to find the required parameters. The outcomes of those approaches and manipulations will naturally imply a conventional crisp soil classification system with some probabilistic statements for different soil types.

Apart from the uncertainty in random nature, there is an uncertainty of fuzziness in the problem of soil engineering classification. Due to the inherent shortfall existing in conventional soil classification methodology, this uncertainty can not be reflected in current **CPT** soil classification Charts. The last part (Chapter 5) of this dissertation, therefore, will begin with discussing the potential use of the theory of fuzzy subset. Theoretically, the fuzzy subset approach will provide us with a powerful tool to describe the general characteristics of soil types. Consequently, an ideal soil classification should take a fuzzy subset system as its carrier. However, how to implement this idea is an issue needed to be fully investigated.

In general, a **CPT** soil Classification problem has its own specific features. The proper combination of the principle of fuzzy subset theory with those specific features is then the key for a successful application of the fuzzy subset theory in this soil classification problem. Such a combination can be achieved only when these specific features are fully understood. Therefore, a detailed discussion on several important characteristics of a **CPT** soil classification problem will be performed. Based upon such an analysis and the basic fuzzy subset theory, a temporary frame of **CPT** fuzzy soil engineering classification will be suggested. Such a frame can be fully expanded and evolved to a mature **CPT** fuzzy soil engineering classification in the future.

Finally also in this part, a new package of **CPT** soil engineering classification based upon the cone tip resistance q_c and friction ratio FR will be suggested and discussed as a summary of the all analysis results obtained in this research.

CHAPTER 1

SIMPLIFICATION OF EXISTING CPT OR PCPT SOIL CLASSIFICATION CHARTS

1.1. Introduction

Cone penetration technology has the outstanding advantage of providing continuous measurements of engineering behaviors of in-situ soils. These measurements include cone tip resistance q_c and cone friction resistance f_s for **CPT** or q_c, f_s and pore pressure u for **PCPT**. A total and comprehensive understanding of site conditions can be achieved quickly if these continuous sounding data are manipulated properly. Unfortunately, this advantage of continuous and visual description of site conditions is not well preserved in the analysis results of site stratigraphy due to the format of **CPT** or **PCPT** soil classification charts currently available. It is known that cone test data are mainly controlled by soil composition and environmental factors (**Douglas and Olsen, 1981**), supposed test equipments and procedures are same. However, current charts have failed to define explicitly the corresponding soil classification and soil in-situ state indices which will greatly improve the efficiency of presenting the analysis results of cones. Therefore, the continuous description of soil conditions along the depth becomes difficult to achieve and the loss of this advantage will eventually prevent users from getting a productive and complete site picture.

This situation also has the potential to obstruct the further development of this technology. The state of art in this technology has shown that the raw data (q_c, f_s, and u) provided by **CPT** or **PCPT** are really not independent from each other in most cases. As mentioned before, people realize that three fundamental factors (equipment and procedure, soil composition, and environmental factors) will control these basic or raw measurements (**Douglas and Olsen, 1981**). Supposing a rigorous standardization on test equipment and test procedure can be followed, the tip resistance q_c, friction resistance f_s and pore pressure u are still affected by both soil composition and environmental factors. So these basic measurements and their normalized results in some cases are really not good candidates to represent parameters on which correlations are established with either compositional and environmental factors or their corresponding engineering properties.

It is clear that there is a merit to find out two independent parameter indices used to represent those two fundamental factors (soil types and soil in-situ states), separately. As an initial effort, these indices can be determined according to current soil classification charts based upon the raw measurements of **CPT** or **PCPT**. If this attempt succeeds, these **CPT** or **PCPT** soil classification charts can be simplified. This chapter will present author's efforts in this issue followed by a brief discussion of the results obtained.

1.2. Observation of Two Tendencies in Soil Classification Charts

Following fact was found in current **CPT** soil classification charts. "All the charts are similar in that sandy soils generally have high cone bearing and low friction ratios whereas, clayey soils generally have low cone bearing and high friction ratios" (**Campanella & Robertson, 1988**). If these charts are carefully studied furthermore, more than that can be concluded. There are two fundamental tendencies with almost orthogonal curve shapes in these charts irrespective of the format of abscissas and ordinates, as exhibited in Figure 1.1, 1.2 and 1.3. Soil type changes in one direction of the tendencies, and the in-situ state of soil (OCR, soil sensitivity) changes in the other direction. These two tendencies happen to be coincidental with two primary factors (components): soil composition and environment. This observation inspired the author to further explore the possibility of finding two independent indices which might represent the two primary components.

Following will be the procedures taken in this attempt. First of all, a curvilineal coordinate system will be established overlapping along the tendencies in each of the charts. Then, this curvilineal coordinate system will be transformed into a cartesian coordinate system by conformal mapping. As a result, two independent indices will be available, one of which can be defined as the soil classification index. The other is the soil in-situ state index.

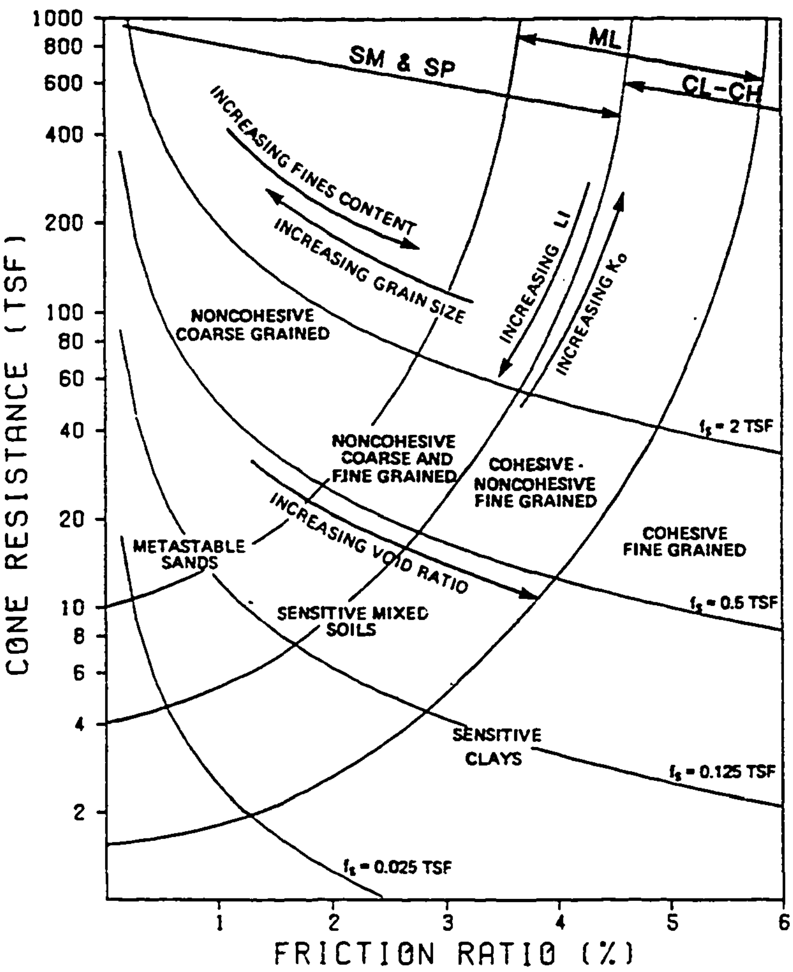


Figure 1.1 Soil Classification Chart for Standard Electric Friction Cone (after B. J. Douglas, et al., 1981)

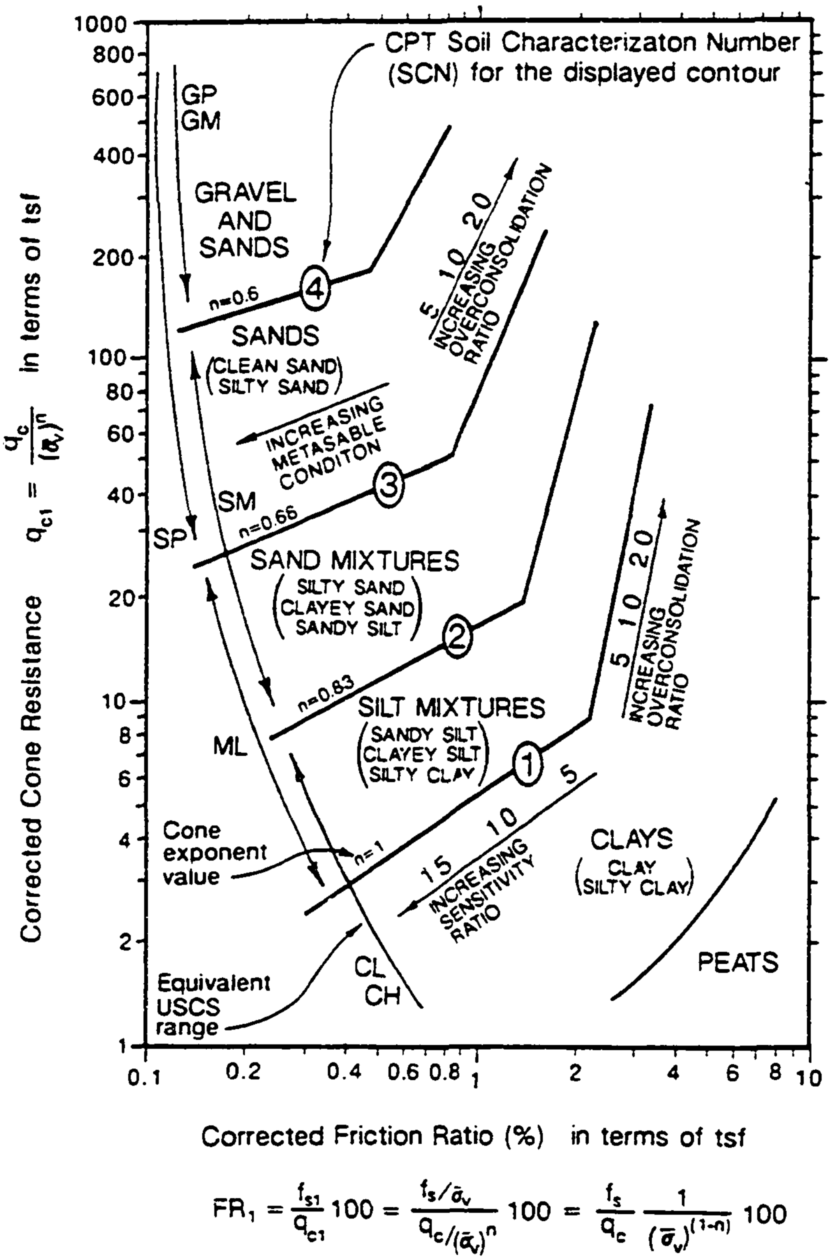
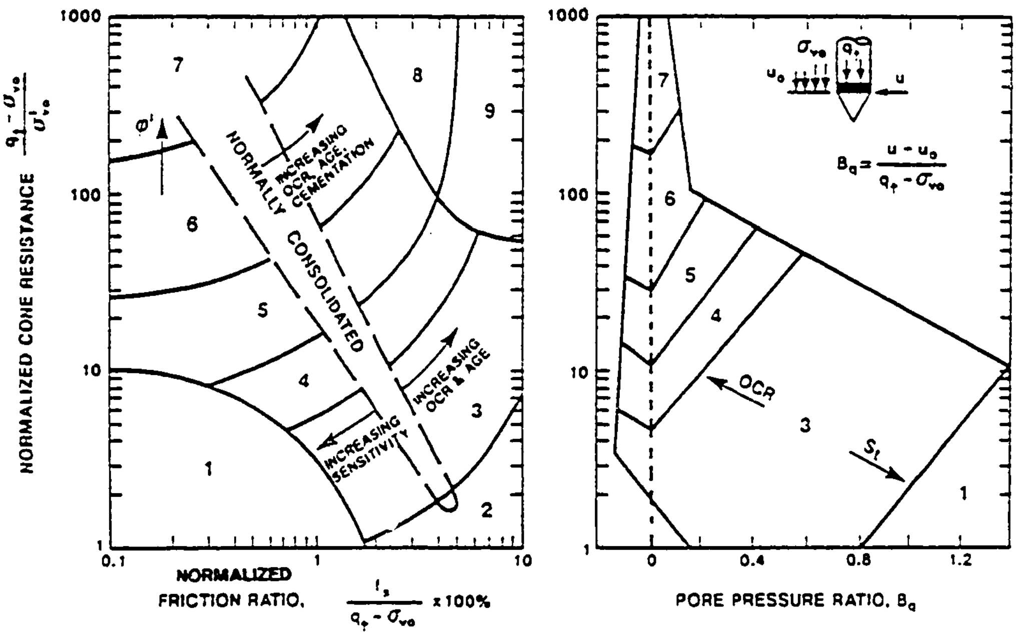


Figure 1.2 Cone Penetration Test (CPT)
Soil Classification Chart
(after R. S. Olsen, et al., 1988)



1. SENSITIVE, FINE GRAINED
2. ORGANIC SOILS – PEATS
3. CLAYS – CLAY TO SILTY CLAY
4. SILT MIXTURES – CLAYEY SILT TO SILTY CLAY
5. SAND MIXTURES – SILTY SAND TO SANDY SILT
6. SANDS – CLEAN SAND TO SILTY SAND
7. GRAVELLY SAND TO SAND
8. VERY STIFF SAND TO CLAYEY* SAND
9. VERY STIFF, FINE GRAINED*

(*) *HEAVILY OVERCONSOLIDATED OR CEMENTED*

Figure 1.3 Proposed Soil Behaviour Type Classification Chart Based on Normalized CPT and CPTU (after P. K. Robertson, 1990)

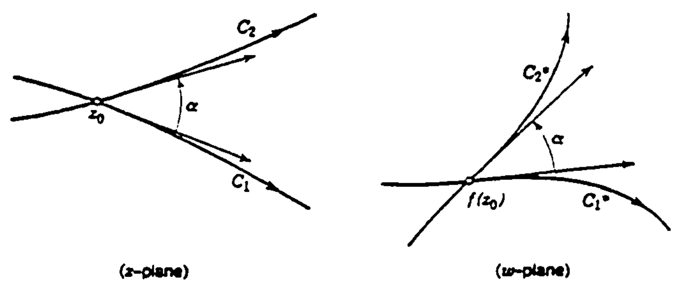


Figure 1.4 Curves C_1 and C_2 and Their Respective Images C_1^* and C_2^* under a Conformal Mapping

1.3. Conformal Mapping

Conformal mapping is a standard method originally used in potential theory to help solving boundary-value problems by transforming a given complicated region into a simpler one. A complex-valued function

$$w = f(z) = u(x,y) + iv(x,y), \quad z = x + iy \qquad (1.1)$$

gives a mapping of its domain with definition in the complex Z-plane onto its range of values in the complex W-plane. Usually, an analytic function (a function f(z) is said to be analytic at a point $z = z_o$ if it can be represented by a power series in powers of $z - z_o$ with radius of convergence $R > 0$) will be used due to its important mapping property of conformality.

Here, a mapping is said to be conformal, i.e., angle-preserving "if it preserves angles between oriented curves in magnitude as well as in sense, that is, if the images of any two intersecting oriented curves, taken with their corresponding orientation, make the same angle of intersection as the curves, both in magnitude and direction. Here the angle between two oriented curves is defined to be the angle α ($0 \leq \alpha \leq \pi$) between their oriented tangents at the point of intersection" (**Kreyszig, 1988**), as Figure 1.4 shows. This property is important in our case since the relative position of any part in a soil classification chart will not be changed in a resulting w plane after a conformal transformation if this chart is treated as a z plane. The orthogonality, if it exists between any two curves in the z plane, will be preserved in the w plane, too. Consequently, the transformed new chart can be expected to have the properties the same as the original chart, i.e., being composed of two soil indices. One represents soil composition. The other reflects the in-situ state of soils.

Following are two quoted theorems from the conformal mapping theory, which provide the tools used later (**Kreyszig, 1988**).

Theorem 1: Every linear fractional transformation (1.1) maps the totality of circles and straight lines in the Z-plane onto the totality of circles and straight lines in the W-plane.

Theorem 2: Three given distinct points z_1, z_2, z_3 can always be mapped onto three prescribed distinct points w_1, w_2, w_3 by one, and only one, linear fractional transformation $w = f(z)$. This mapping is given implicitly by the equation

$$\frac{w - w_1}{w - w_3} \; \frac{w_2 - w_3}{w_2 - w_1} = \frac{z - z_1}{z - z_3} \; \frac{z_2 - z_3}{z_2 - z_1} \qquad (1.2)$$

A conformal transformation $w = f(z)$ is determined as the result of applying these theorems to our case. This function is assumed to satisfy the following three pairs of corresponding points on a Z plane and the corresponding W plane:

$z_1 = 0.63 + i$, $\quad z_2 = 1.34$, $\quad z_3 = 3 - 0.73\, i$

$w_1 = -5$, $\quad w_2 = 0$, $\quad w_3 = 5$

The choice of these z and w values is the result of a "trial and error" procedure due to the empirical nature of this approach. The only guideline here is that the result obtained can serve our purpose. Substituting all these values into Equation (1.2) and simplifying it , we get (the derivation can be found in Appendix A)

$$u = \frac{(a_1x - a_2y + b_1)(c_1x - c_2y + d_1)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} + \frac{(a_2x + a_1y + b_2)(c_2x + c_1y + d_2)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} \qquad (1.3)$$

$$v = \frac{(c_1x - c_2y + d_1)(a_2x + a_1y + b_2)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} - \frac{(a_1x - a_2y + b_1)(c_2x + c_1y + d_2)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} \qquad (1.4)$$

here, $a_1 = -11.345$, $a_2 = -3.795$, $b_1 = 15.202$, $b_2 = 5.085$,
$c_1 = -0.269$, $c_2 = -0.759$, $d_1 = -2.960$, $d_2 = 2.477$.

These formulas will be used as the basic conformal transformation tools in the subsequent research.

1.4. Transformation of Soil Classification Charts

Literature review has shown that there are many soil engineering classification charts of cone technology (**Schmertmann, 1978; Douglas & Olsen, 1981; Jones and Rust, 1982; Senneset and Janbu, 1984; Tumay 1985; Robertson et al, 1986; Olsen and Malone, 1988; Campanella and Robertson, 1988; Senneset et al, 1989; Robertson, 1990; Cheng-hou et al, 1990**). Among them are B. J. Douglas et al's chart (**1981**), R. S. Olsen et al's chart (**1988**), and P. K. Robertson's soil classification charts (**1990**), as shown before in Figure 1.1, 1.2, and 1.3. Olsen's chart and Robertson's charts are most popular currently and Douglas' chart is a good landmark to examine the improvement obtained by normalization. Therefore, they are to be taken as examples in this research to illustrate the conformal mapping approach.

As Figure 1.1 shows, there is no normalization adopted in Douglas' soil classification chart. The horizontal axis in that chart is the fraction ratio $FR = f_s/q_c$ (%). The vertical axis is the tip resistance q_c (tsf). The corresponding curvilineal coordinates, which reflect the changes of soil types and soil in-situ states, are shown in Figure 1.5.

In Olsen's chart as shown in Figure 1.2, the CPT data have been normalized with respect to effective vertical stress σ'_v. The abscissa is the corrected friction ratio $FR_1 = f_{s1}/q_{c1} = FR/(\sigma'_v)^{(1-n)}$ (%). Here, the corrected CPT sleeve friction resistance f_{s1}, in terms of tsf, is equal to f_s/σ'_v and n is the exponent value. The ordinate is the corrected cone tip resistance $q_{c1} = q_c/(\sigma'_v)^n$, also in terms of tsf. The corresponding curvilineal coordinates are shown in Figure 1.7.

There are two charts in Robertson's soil classification system, as presented in Figure 1.3. Chart (1) is based upon tip resistance q_c and the friction resistance f_s. Chart (2) is from tip resistance q_c and pore pressure u. q_c should first be corrected to a total cone resistance, q_t, using the following expression:

$$q_t = q_c + (1 - a)\,u \tag{1.5}$$

where "u" is pore pressure measured between the cone tip and the friction sleeve and "a" is net area ratio. Then the vertical axes in both charts are in term of $Q_t = (q_t - \sigma_{v0})/\sigma'_{v0}$ and the horizontal axis in chart (1) is $F_R = f_s/(q_t - \sigma_{v0})$ 100% and the one in chart (2) is $B_q = (u - u_0)/(q_t - \sigma_{v0})$. Here, u_0 is equilibrium pore pressure, σ_{v0} is total overburden confining stress, and σ'_{v0} is effective confining stress. Figure 1.9 and 1.11 show the corresponding curvilineal coordinates.

Several assumptions have to be taken here in order to make the conformal mapping approach more simple and meaningful. First, the changes of the soil composition and environmental factors can be represented by the two basic tendencies in these charts. This assumption is acceptable due to the fact we observed in all current soil classification charts of cone technology.

The second assumption is that the curves of the curvilineal coordinates in soil classification charts are supposed to be circle arcs so that the corresponding locus function of each circle arc, in terms of rectangular coordinates, can be easily determined by three pairs of rectangular coordinates. The most important benefit of this assumption is obvious here, i.e., the circle arcs can be easily transformed to straight lines by a linear fractional transformation provided by Equation 1.2.

The third assumption is that the interaction between soil composition and environmental factors is negligible. Thus, the soil composition and environmental factors can be treated as independent so that the curves of the curvilineal coordinates can be assumed to be orthogonal to each other. Consequently, the curvilineal coordinates can, under the second assumption, be easily transformed to a rectangular coordinate system. The validation of all these assumptions will be checked later by the results obtained.

Following are the results from a "trial and error" procedure, where the intermediate variable u and v are defined by Equation (1.3) and (1.4), separately. The complex plane Z in Douglas' chart has the following form:

$$x = 0.1539\,FR + 0.8870\log q_c - 3.35 \qquad (1.6)$$

$$y = -0.2957\,FR + 0.4617\log q_c - 0.37 \qquad (1.7)$$

here, FR is in % and q_c is in tsf. The final index U = - u and V = v + 10. Figure 1.5 is the soil classification chart (FR --- log q_c plane) with the curvilineal coordinates U and V. Figure 1.6 shows the corresponding U --- V plane with the transformed Douglas' chart, including the boundary curves of the soil classification.

Similarly, suppose the complex plane in Olsen's chart has the following form:

$$z = x + iy = \log FR_1 + i\,(\log q_{c1} + 1.6) \qquad (1.8)$$

or

$$z = (\log f_{s1} - \log q_{c1}) + i\,(\log q_{c1} + 1.6) \qquad (1.9)$$

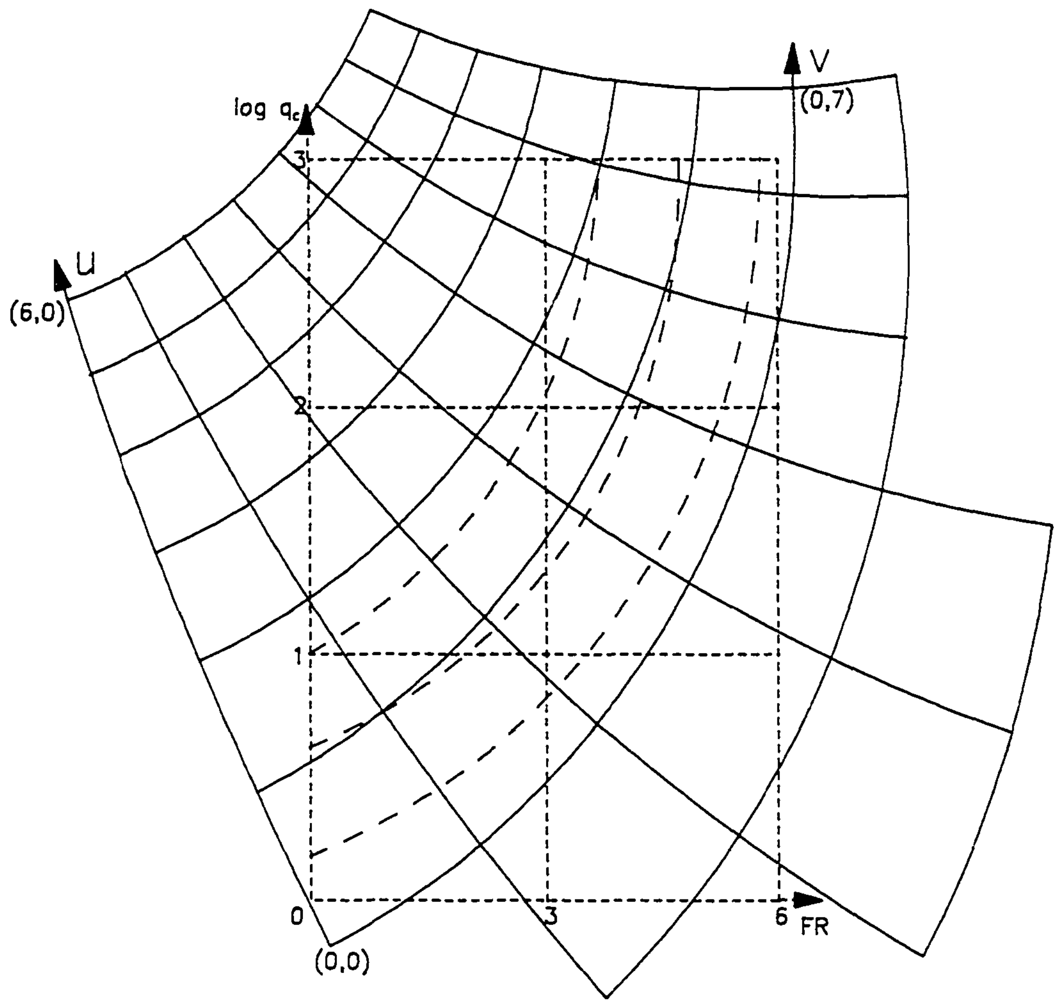


Figure 1.5 Douglas Chart (FR --- log q_c Plane) with Curvilineal Coordinates U and V

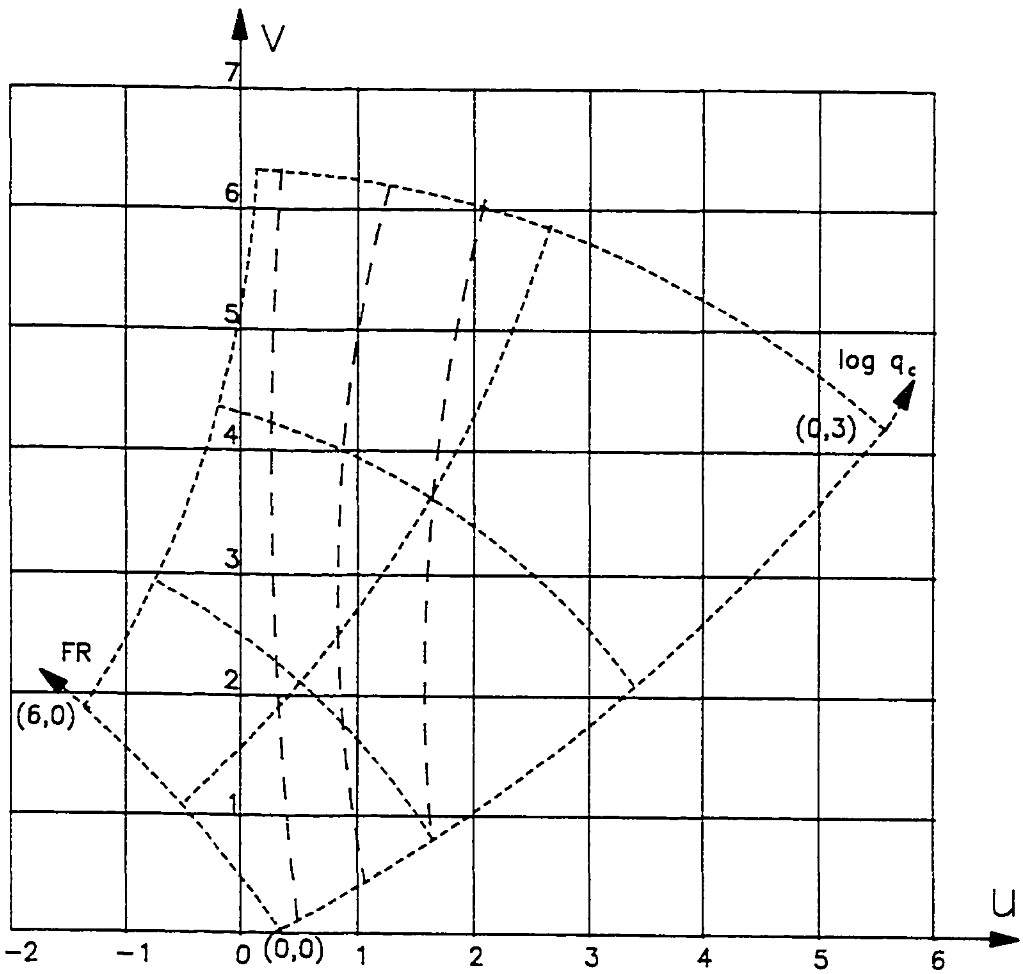


Figure 1.6 U --- V Plane with Transformed Douglas Chart

here, FR_1 is in % and q_{c1} is in tsf. The final index U = - u and V = v. Figure 1.7 gives out the soil classification chart (log FR_1 --- log q_{c1} plane) with the curvilineal coordinates U and V. Figure 1.8 presents the corresponding U --- V plane with the transformed Olsen's chart, including the boundary curves of the soil classification.

For Robertson's charts, the complex plane in chart (1) will take the form:

$$x = 0.7009\,(\log F_R + 1) + 0.7132\log Q_t - 1.82 \qquad (1.10)$$

$$y = -0.7132\,(\log F_R + 1) + 0.7009\log Q_t + 0.97 \qquad (1.11)$$

Here F_R is in % and Q_t is in MPa. The final index U = -(u + 1) and V = v. Figure 1.9 displays the soil classification chart (1) (log F_R --- log Q_t plane) with the curvilineal coordinates U and V. Figure 1.10 shows the corresponding U --- V plane with the transformed Robertson's chart (1), including the boundary curves of the soil classification.

The Z plane in chart (2) will be, for $B_q \geq 0$,

$$x = 0.5918B_q + 0.9348\log Q_t - 1.95 \qquad (1.12)$$

$$y = -1.5580B_q + 0.3551\log Q_t + 1.62 \qquad (1.13)$$

and, for $B_q < 0$,

$$x = 0.6243B_q + 0.9272\log Q_t - 3.1746 \qquad (1.14)$$

$$y = -1.5453B_q + 0.3746\log Q_t + 1.9 \qquad (1.15)$$

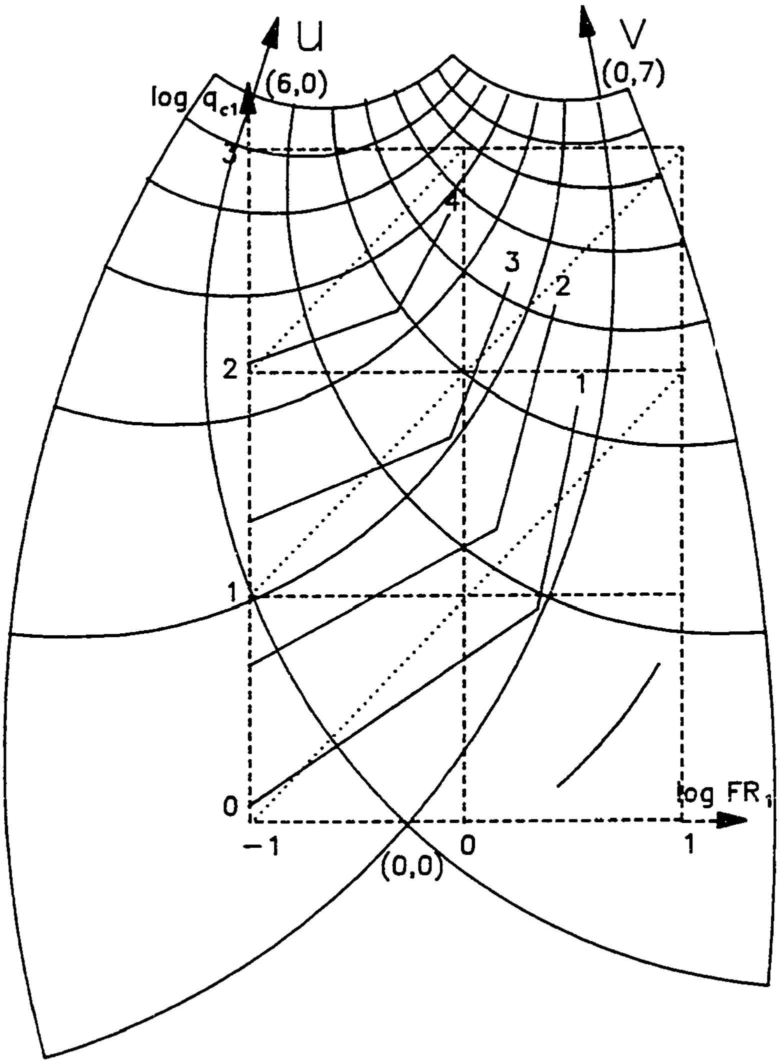


Figure 1.7 Olsen Chart (log FR_1———log q_{c1}) with Curvilineal Coordinates U and V.

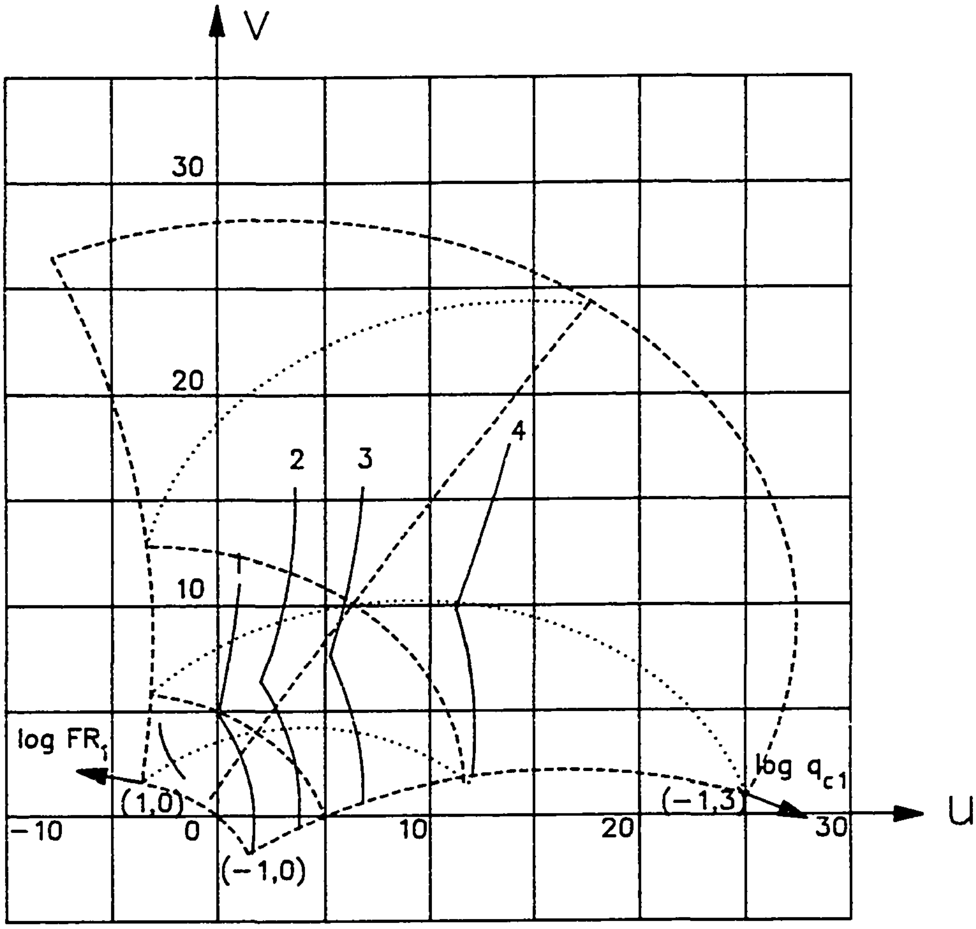


Figure. 1.8 U —— V Plane with Transformed Olsen Chart.

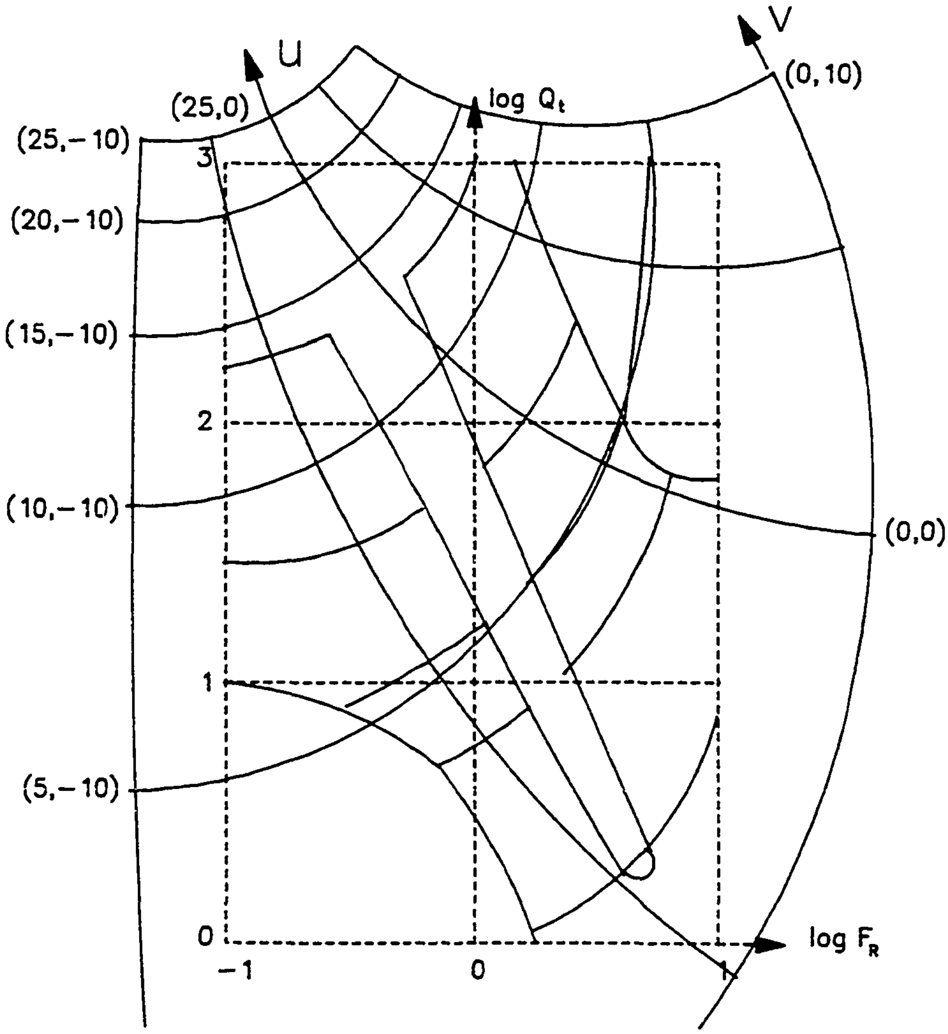


Figure 1.9 Robertson Chart (1) with Curvilineal Coordinates U & V.

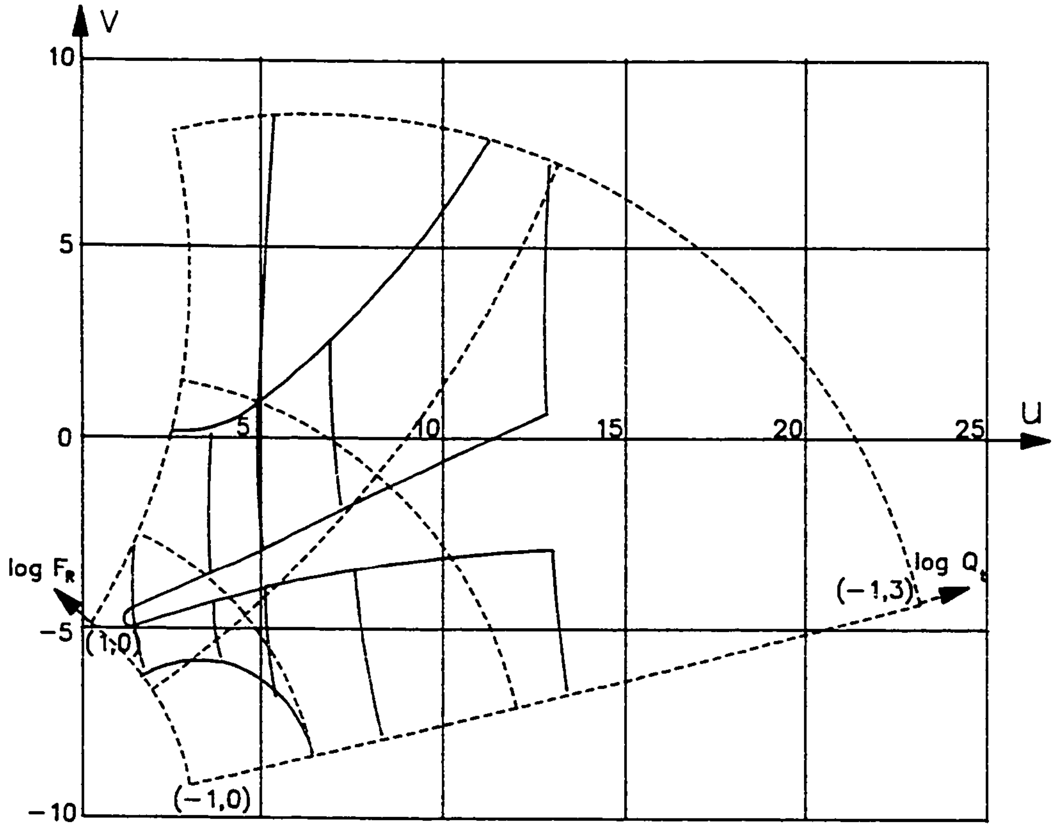


Figure 1.10 U --- V Plane with Transformed Robertson Chart (1)

The final index $U = -u$ and $V = v + 12$. Its B_q --- $\log Q_t$ plane with the curvilineal coordinates U and V is shown in Figure 1.11 and the U --- V plane with the transformed Robertson's chart (2) is exhibited in Figure 1.12, including the boundary curves of the soil classification. Obviously, the index U where $B_q > 0$ should be treated differently from the U where $B_q < 0$.

Since soil type, as already observed, changes along the U axis and soil in-situ state changes along the V axis in all these transformed charts, it is reasonable to use U to identify different types of soils and V to distinguish different in-situ states of soils. Therefore, U has been named as "soil classification index" and V has been called the "soil in-situ state index". However, whether this name is proper in the case of Robertson's chart (2) is not clear for the time being.

It is apparent from all these figures that the soil classification index U in each simplified (or transformed) soil classification system can be used as good as each corresponding original soil classification chart. This fact indicates that the assumptions made during the analysis are reasonable and acceptable since all the original classification charts are the culmination of experiences to the date. The biggest benefit of using these simplified soil classification systems is that the advantage of continuous and visual description of site situation in raw cone data is well preserved in the analysis results of site stratigraphy. Consequently, the efficiency of expressing the interpretation of cone data is greatly improved, as the example in Figure 1.13 and 1.14 shows. Figure 1.13 presents

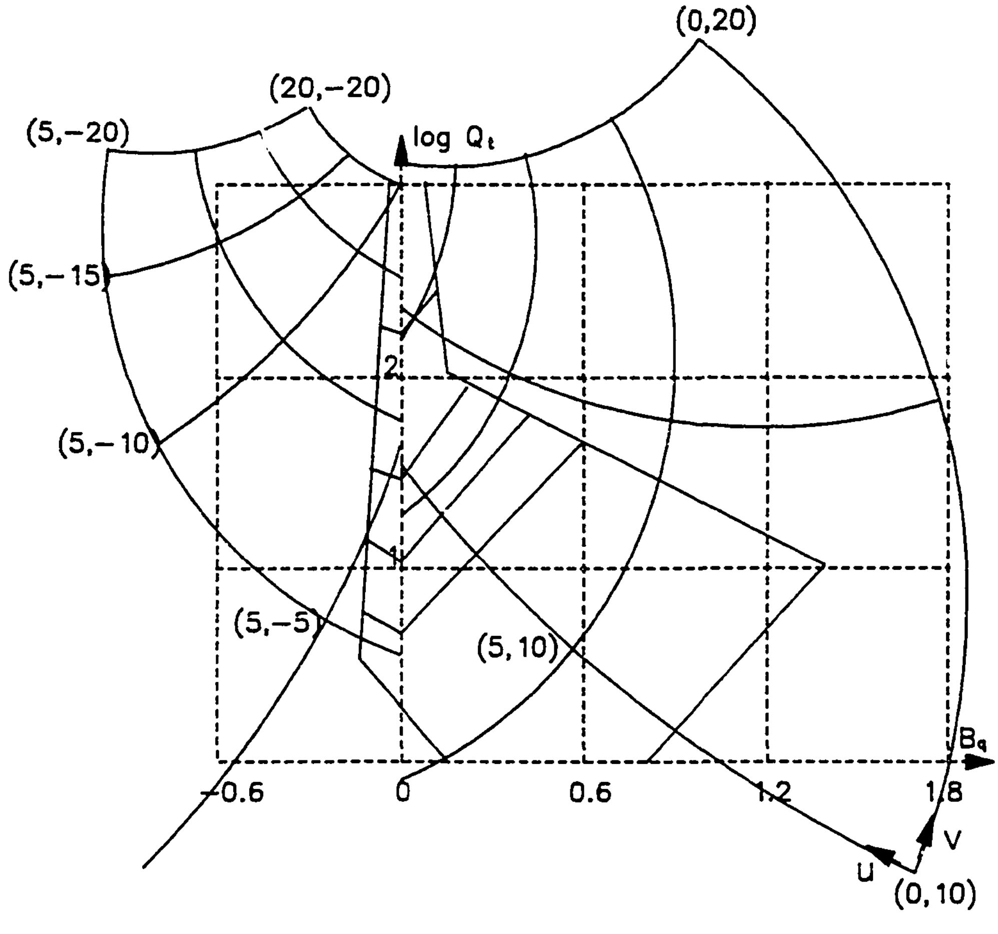


Figure 1.11 Robertson Chart (2) with Curvilineal Coordinates U & V.

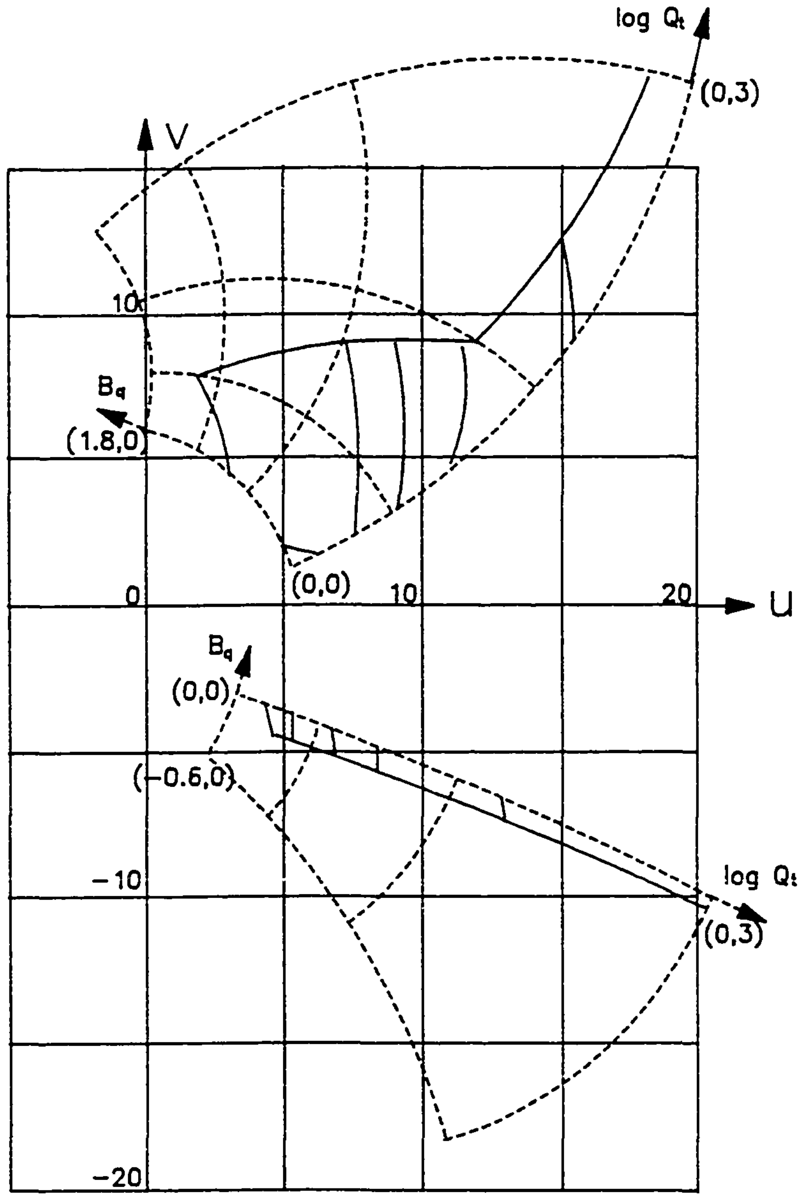


Figure 1.12 U ——— V Plane with Transformed Robertson Chart (2).

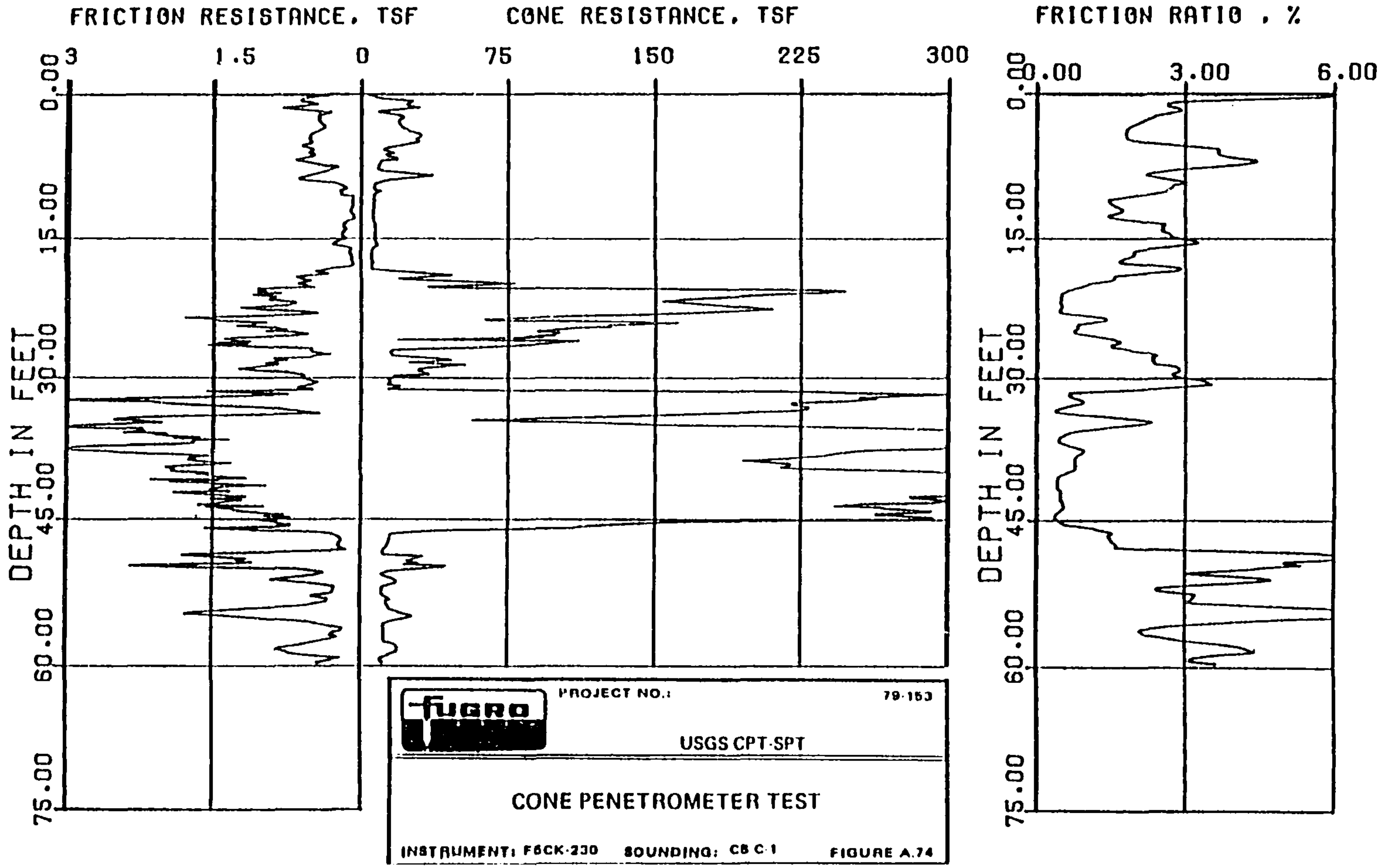


Figure 1.13 A Profile of Raw CPT Data from San Jose Site.
(from USGS Open-File Report No. 81-284, 1980)

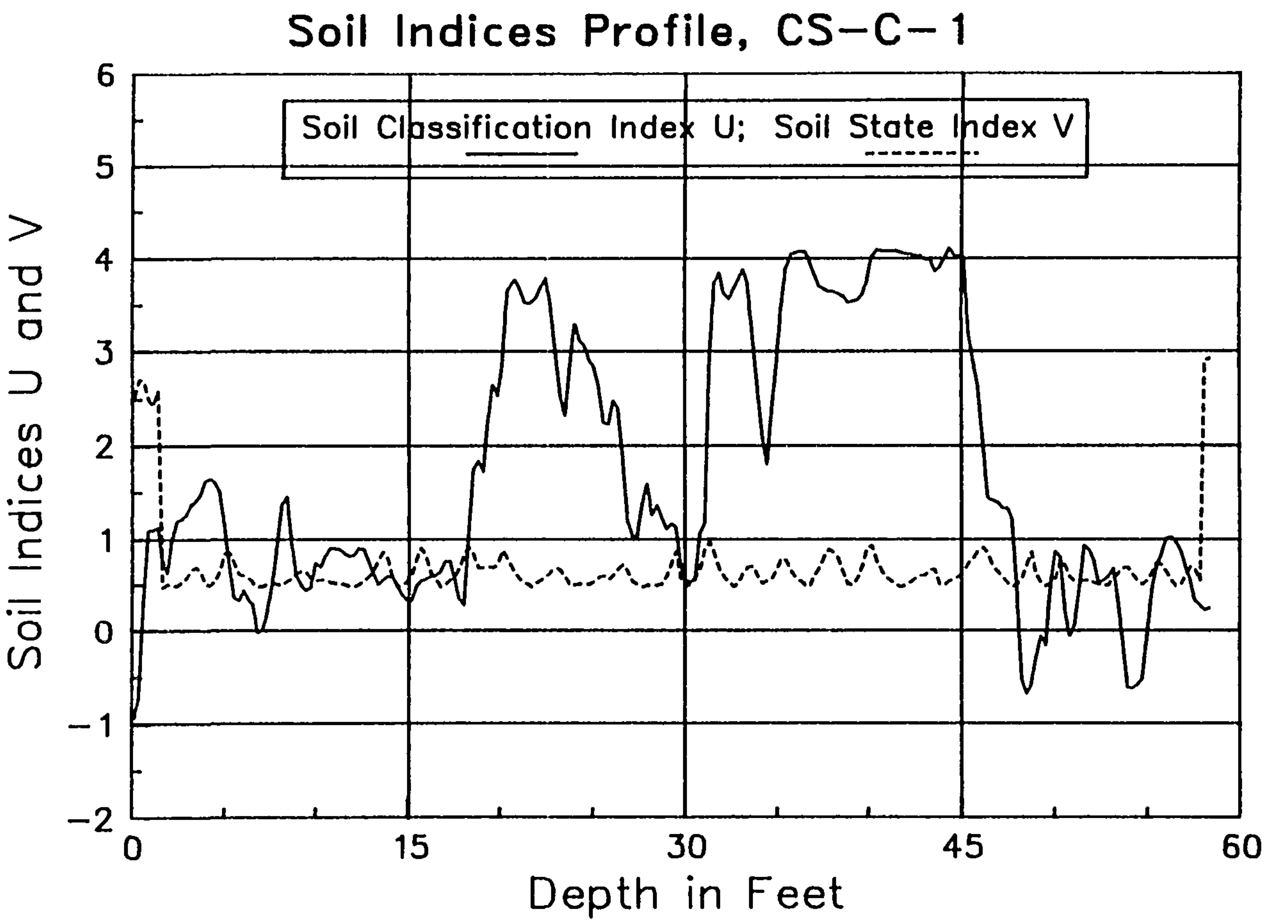


Figure 1.14 The Corresponding U & V Profile of Figure 1.13, Based upon the Transformed Douglas Chart.

a profile of raw **CPT** data (**USGS Open-file Report No. 81-284, 1980**). Figure 1.14 displays the corresponding U and V profiles based upon the transformed Douglas' chart. If appropriate soil classification criteria are given in Figure 1.14, the whole picture of in-situ soil types will be very clear.

1.5. Verification from Some In-situ Data

CPT data from four sites have been used to check the two one-dimensional soil classification systems simplified from Douglas' chart and Olsen's chart, separately. Two sites of them are San Diego (9 CPT soundings) and Salinas (14 CPT soundings), California. The detailed description of them can be found in reference (**USGS Open-file Report No. 81-284,1980**). The electric friction cone used in these tests is the Fugro tension type cone with a cylindrical friction sleeve of 150 cm^2 surface area, capped with a 60-degree apex angle conical tip of 10 cm^2 projected surface area. More detailed information on the data is available in the relevant reference. The other two sites, which are from personal resource, are two projects within the Theodore area in Alabama. The electric friction cone used in these tests is a 43.7 mm nominal diameter Fugro cone penetrometer (cross-sectional area of 15 cm^2), with a friction sleeve area of 200 cm^2 and a cone apex angle of 60°, being a substraction type. The soundings (6 and 9 separately) were performed using Louisiana State University Research Vehicle for Geotechnical Insitu Testing and Support (REVEGITS). The testing procedure was in general accordance with the procedure recommended by the **ASTM**. The corresponding boring profiles are provided by Southern Earth Science, Inc.

Following linear regression analysis results from Lima and Tumay (1991) have been used in order to make all the data comparable to each other:

$$q_{c,(10cm^2)} = -0.073 + 1.055\,q_{c,(15cm^2)} \qquad (1.16)$$

$$R_{f,(10cm^2)} = 0.812 + 0.931\,R_{f,(15cm^2)} \qquad (1.17)$$

Here q_c is in MN/m^2 and R_f is friction ratio, FR. Consequently, the succeeding analysis will be based on the data from the cone with 10 cm^2 cross-sectional area or the equivalent ones.

Since soil classification based upon CPT is a behavior type, how to describe the behavior of a soil type becomes an essential problem to solve. It appears that a concept of soil behavior unit in cone penetration is necessary so that soil behavior can be characterized in quantity. Only under such a condition, can a good correlation between soil types and soil behaviors be found out. It can be predicted that this behavior unit property should contain the information of not only the values of q_c and f_s but also the spatial distributions of them. The easiest way to implement this idea is to find soil layers with approximately constant q_c and FR values or similar variations. This is because a layer in this context implies the spatial distribution.

Therefore, soil behavior units in cone penetration are empirically determined for the time being as follows. First, the top and bottom elevations of each soil layer are determined by examining given boring and sounding logs, sounding data being given more weight.

Soil types are decided according to the boring logs. Second, if necessary, each layer of soils is further divided into sublayers of same soil type based upon the values of q_c and FR or the variation patterns of them along the depth. Third, the average value of the soil classification index, U, is calculated for each layer according to the corresponding averages of q_c and FR. This average U value is then taken as the soil behavior unit of the corresponding soil layer.

Table 1.1. Average Unit Weight for Different Soils

Soil Type	Average Soil Unit Weight (tcf)		
	γ_s	γ_d	γ'
Sands & Gravels	0.06714	0.05933	0.03591
Silts & Clays	0.05464	0.03747	0.02342
Organic Silts & Clays	0.04840	0.03122	0.01717

Table 1.2. Exponent Value n for Different Soils

Soil Type	Gravel	Sand	Silt	Clay
n	0.6	0.66	0.83	1.0

In order to get normalized soil behavior for each layer with respect to effective vertical stress in the way suggested by R.S. Olsen (**1988**), effective vertical stresses have to be calculated. Table 1.1 gives some average values of soil unit weight, which are modified after R. D. Holtz (1981), and is used whenever real data is not available. If this is the case, γ_{sat} will empirically be taken as the unit weight for clay above the ground water level. Also, the exponent values (n) are modified after R.S. Olsen, as Table 1.2 shows. The names of soil types in this table indicate the basic material in a layer of soils. For example, both silty sand and clayey sand will take 0.66 as their n values. The justification for doing these might be questionable but can be checked through final analysis results. One plausible argument is that the eventual soil classification system should be statistical in nature so that the error caused by these simplifications might be insignificant.

The outcomes of treatments on the in-situ data are presented from Figure 1.15 to Figure 1.24. The results in the former five charts are calculated according to the simplified Douglas' soil classification system which is based upon FR and log q_c. The results in the latter five charts however are according to the simplified Olsen's soil classification system which is based upon log FR_1 and log q_{c1}.

Figure 1.15 is the scatter of soil types in the conformably mapped U --- V plane. The U and V values here are layer based, i.e., they represent soil behavior units. This meaning will be used in subsequent sections in this chapter without explanation if no ambiguity occurs. It can be seen from Figure i.15 that the data do not confirm Douglas' soil classification chart well, which reflects the non-normalized behavior of soil types. Douglas' chart is believed to be correct "only for a vertical effective stress of 1 tsf (i.e., based on penetrations of 15 to 25 feet (5 to 8 meters))" (**Olsen and Farr, 1988**). The data used in this research is well out of this range, which can further be

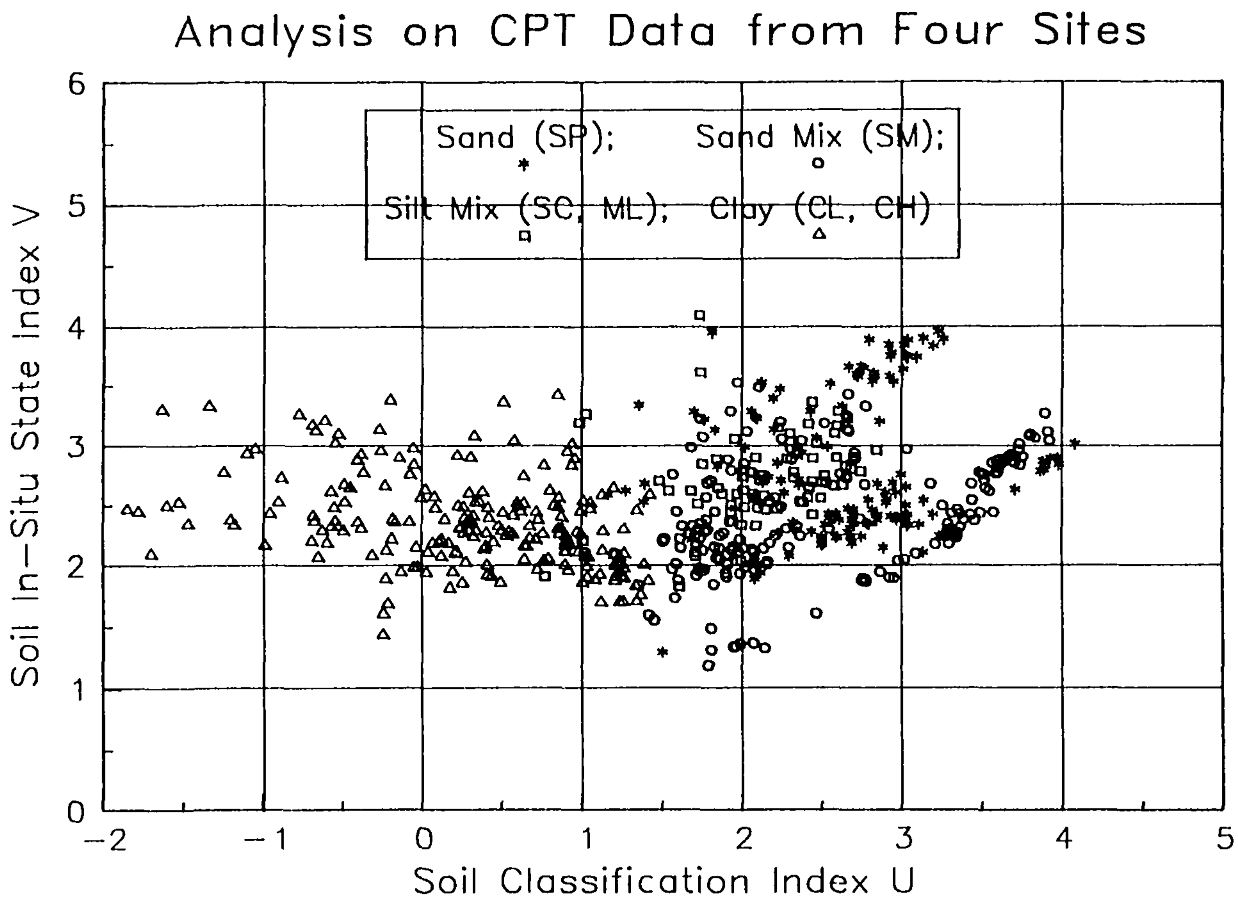


Figure 1.15 Scatter of Soil Types in U --- V Plane
Calculated Based on FR & log q_c.

verified in Figure 1.17. Figure 1.16 shows the scatter relation between thicknesses of layers and soil classification index, U. It can be said from this figure that there is no dependence between thicknesses of layers and soil classification index, U.

Figure 1.17 is the scatter relation between depths of layers and soil classification index, U. It can be seen from this chart that for the same types of soil, the values of classification index, U, will generally increase as their depths of positions increase. This phenomenon is true for the data from the same testing sites. Therefore, this chart has confirmed the claim that q_c and FR should be normalized with respect to the depth (equivalent to effective vertical stress) in order to get the normalized behavior of soils.

Figure 1.18 and Figure 1.19 are the scatters indicating the possible relations between relative STD (standard deviation) of tip resistance (**RSTDTR**) or friction ratio (**RSTDFR**) and soil classification index, U. Here, **RSTDTR** and **RSTDFR** are defined as (STD of q_c)/(average of q_c) and (STD of FR)/(average of FR), separately. All these quantities are layer based. Although the data from each individual site has shown some degrees of dependencies between **RSTDTR** and U or **RSTDFR** and U, these tendencies are not well defined and will disappear if all data are put together. Notice that the values of **RSTDTR** and **RSTDFR** are generally smaller than one (1), which reflects the error caused by the empirical layering procedure adopted in this chapter.

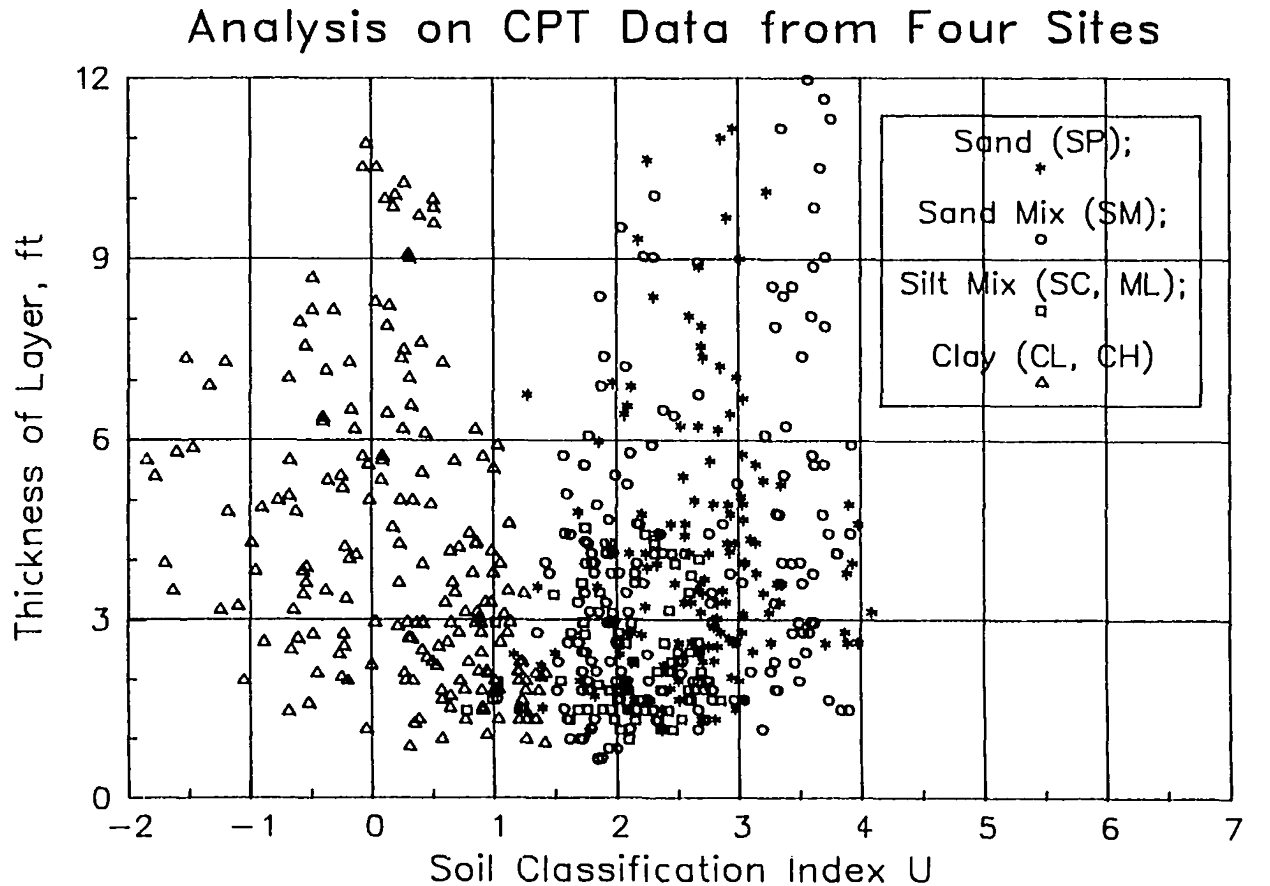


Figure 1.16 Scatter Indicating the Relation between Thinckness of Layers and Classification Index U, Based upon FR and log q_c.

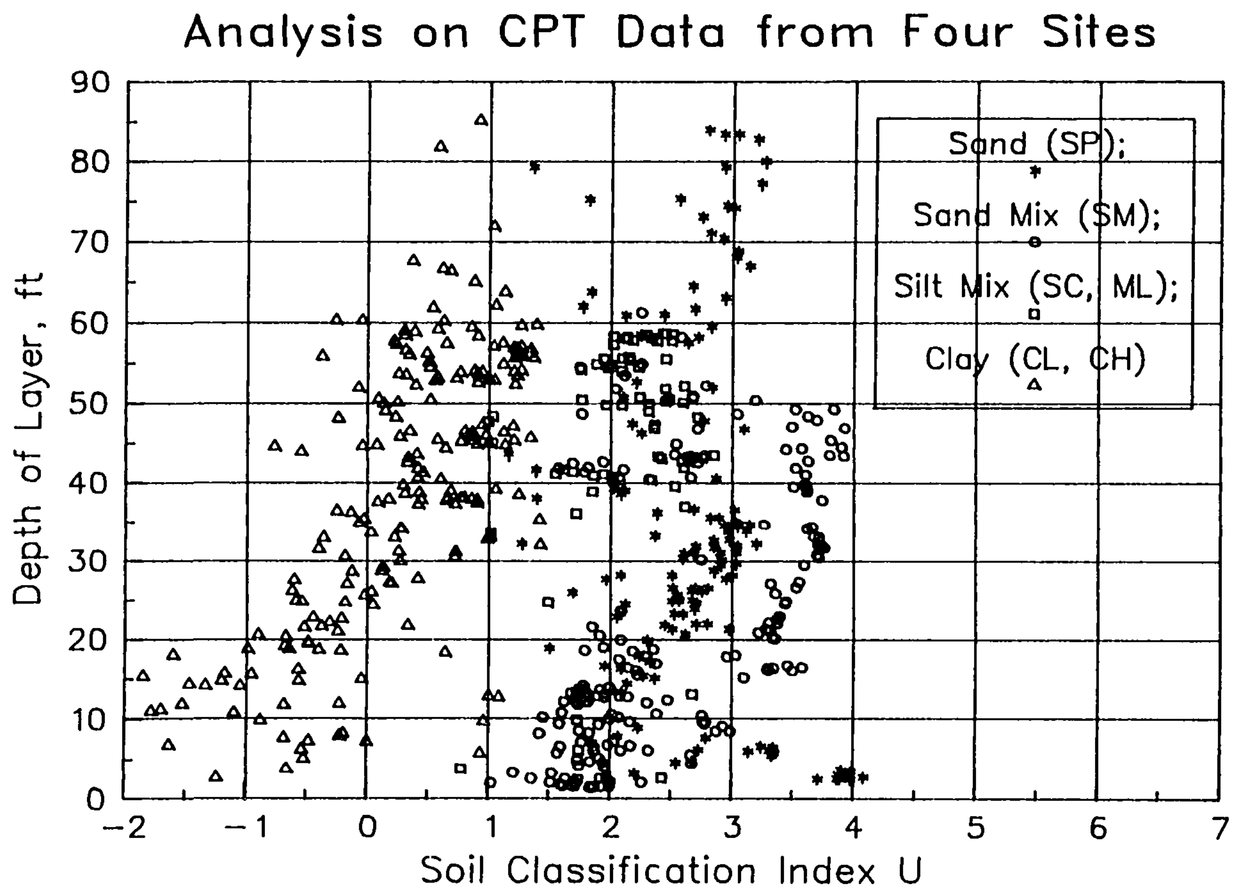


Figure 1.17 Scatter Indicating the Relation between Depth of Layers and Classification Index U, Based upon FR and log q_c.

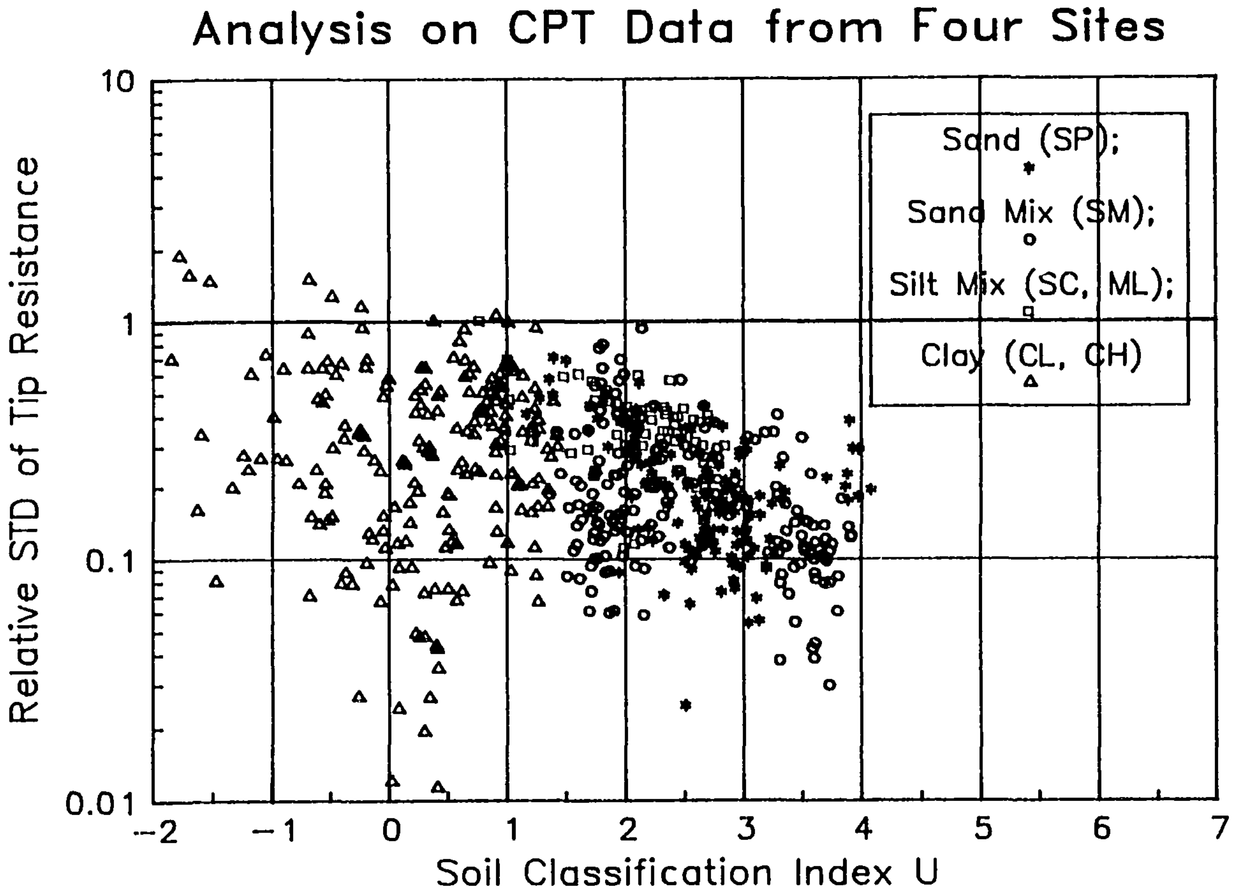


Figure 1.18 Scatter Indicating the Relation between Relative STD of Tip Resistance and Classification Index U, Based upon FR and log q_c.

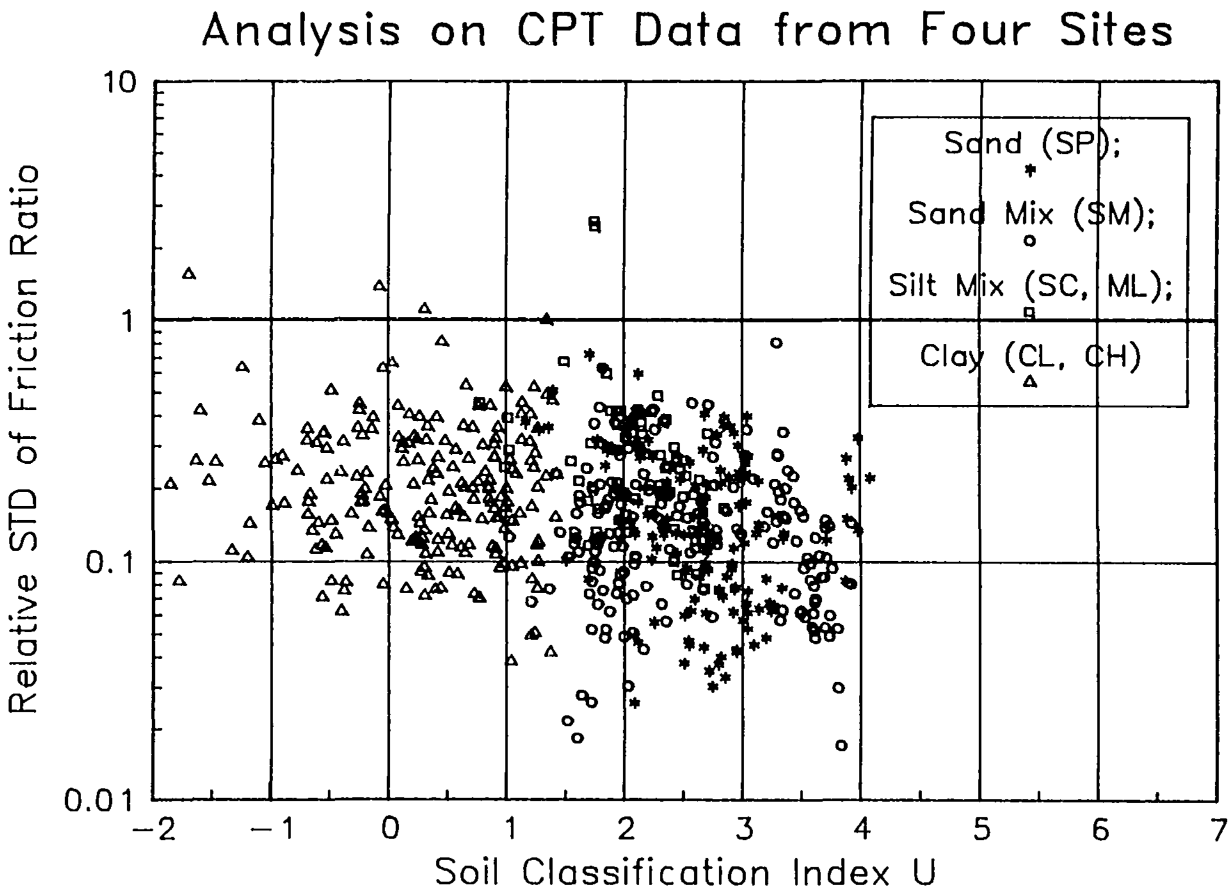


Figure 1.19 Scatter Indicating the Relation between Relative STD of Friction Ratio and Classification Index U, Based upon FR and log q_c.

Figure 1.20 is the scatter of soil types in the transformed U --- V plane calculated based upon log FR_1 and log q_{c1}. The boundary values of different soil types in this figure are in general accordance with Olsen's soil classification chart. Other results from the similar data treatment, shown in Figure 1.21 and 1.22, have indicated that the soil classification index, U, of the simplified Olsen's classification system is independent of thicknesses and depths of soil layers so that Olsen's soil classification chart is indeed based upon the normalized behavior of soils with respect to vertical confining stresses. Also, the error caused by the layering procedure used in this chapter is not very large, as shown in Figure 1.23 and 1.24. Here, the relative STDs of normalized tip resistance and friction ratio are layer based and defined as (STD of q_{c1})/(average of q_{c1}) and (STD of FR_1)/(average of FR_1), separately.

One important fact should be noticed here. Different soil types do overlap with one another in these soil classification systems of soil-behavior type although only the data of four sites have been used in this analysis. It might be predicted that only a statistical certainty can be achieved in these systems. This topic will be thoroughly studied in Chapter 3.

1.6. Some Comments

The basic idea of this chapter was inspired by the observation on current **CPT** or **PCPT** soil classification charts available and this idea has been implemented through the conformal mapping theory in complex analysis. The mathematical tool can be

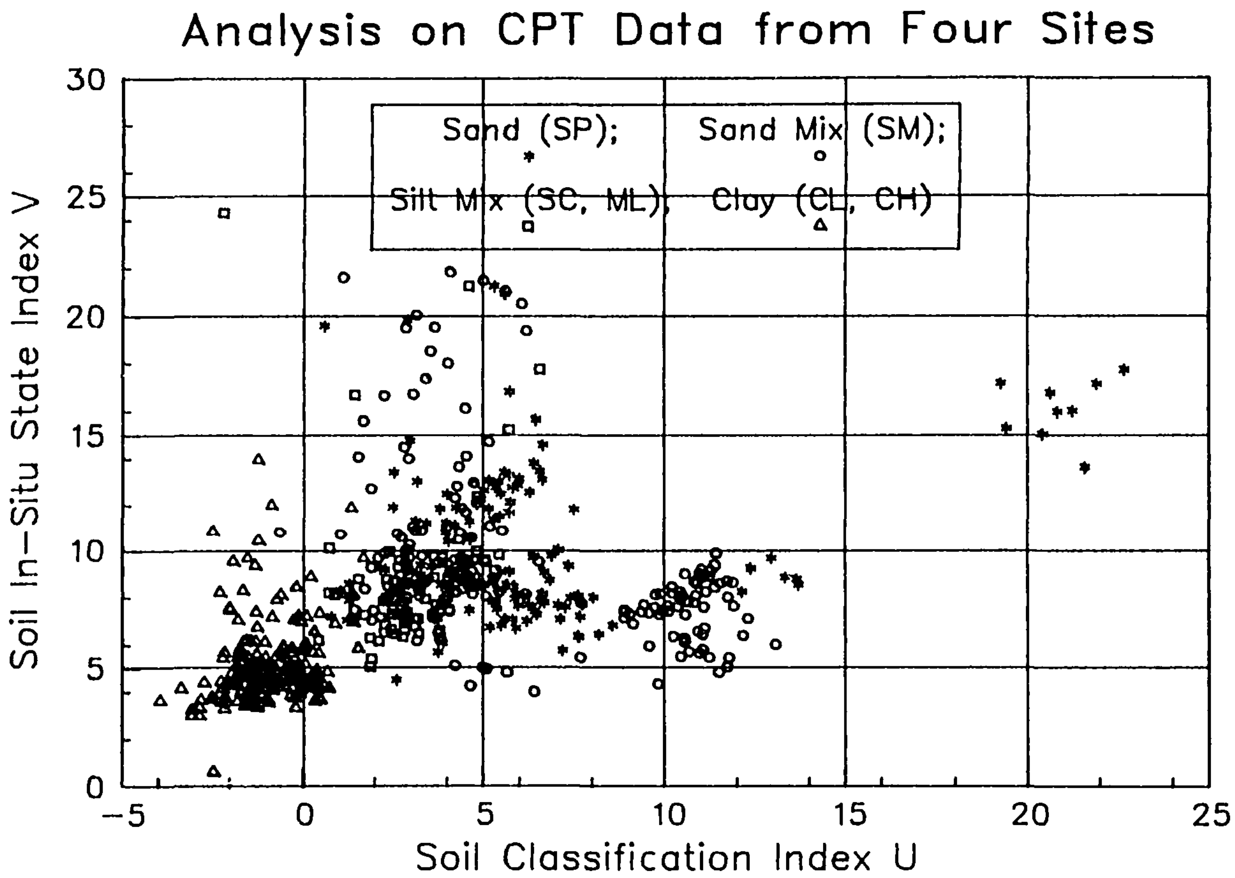


Figure 1.20 Scatter of Soil Types in U --- V Plane Calculated from log FR_1 & log q_{c1}.

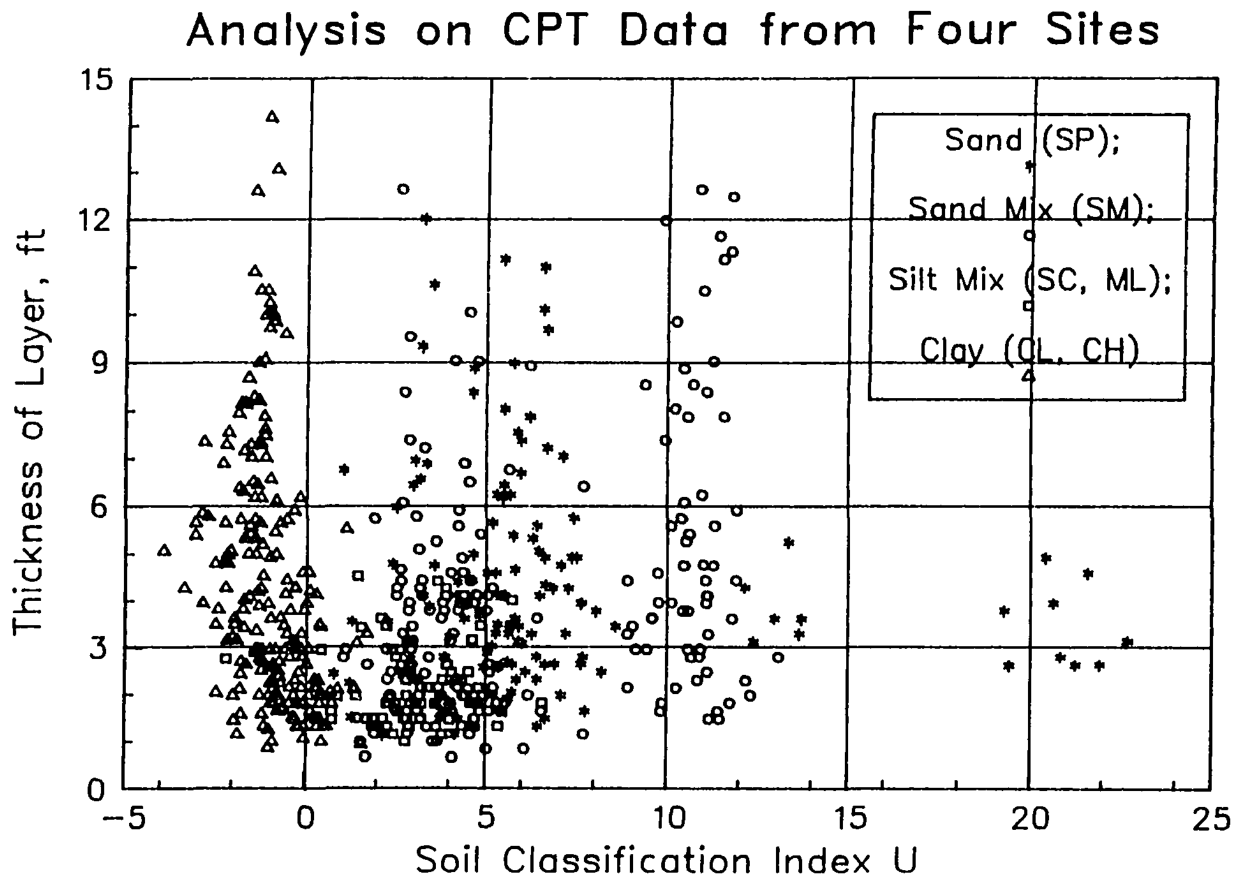


Figure 1.21 Scatter Indicating the Relation between Thickness of Layers and Classification Index U, Based upon log FR_1 and log q_{c1}.

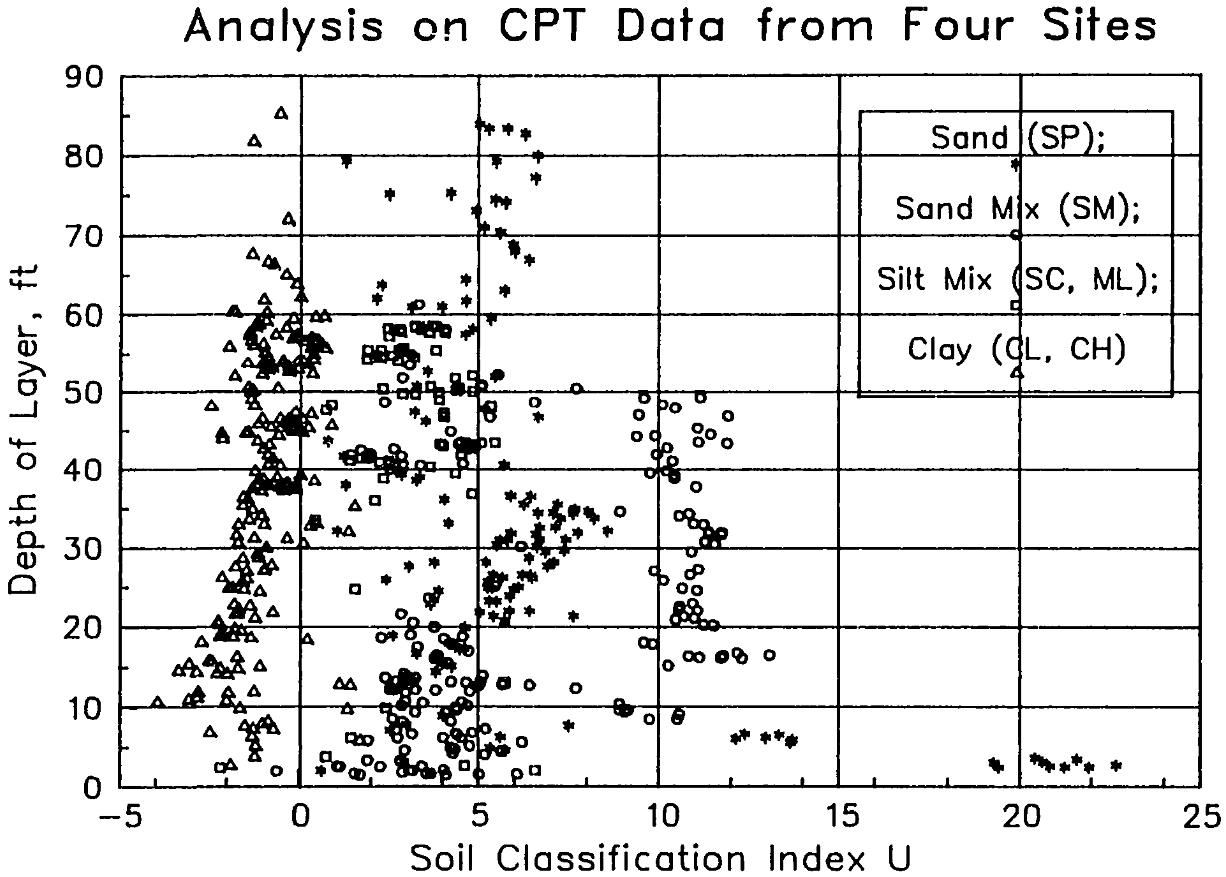


Figure 1.22 Scatter Indicating the Relation between Depth of Layers and Classification Index U, Based upon log FR_1 and log q_{c1} .

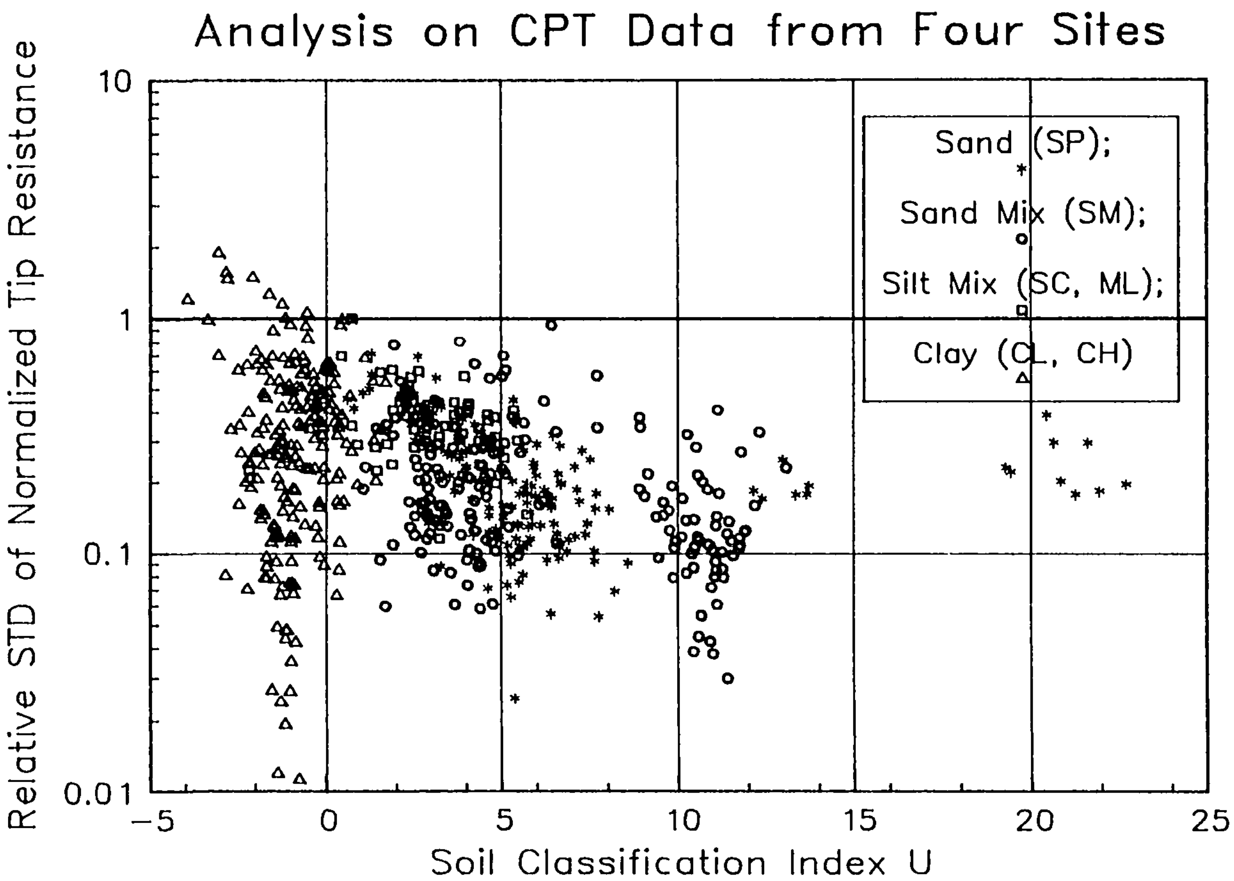


Figure 1.23 Scatter Indicating the Relation between Relative STD of Normalized Tip Resistance & Classification Index U, Based upon log FR_1 & log q_{c1} .

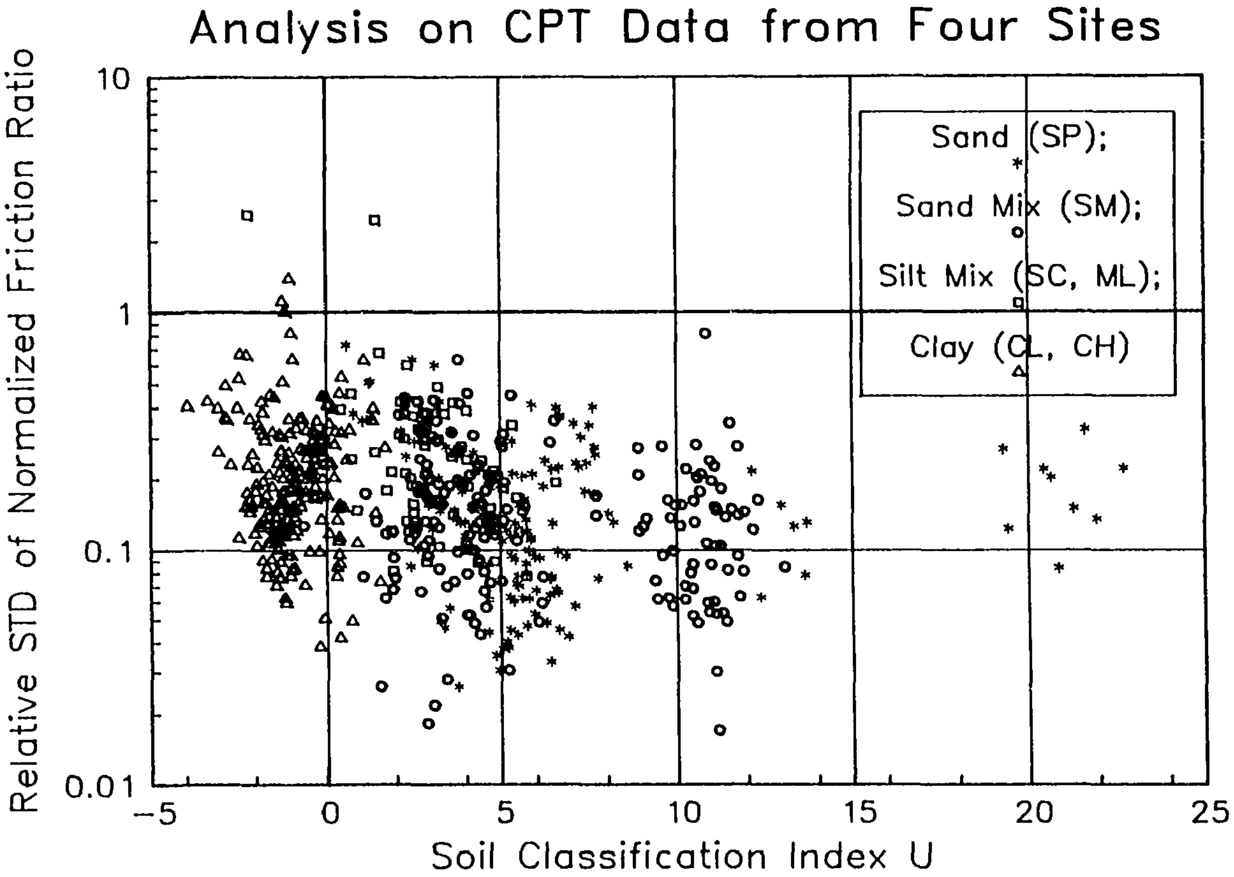


Figure 1.24 Scatter Indicating the Relation between Relative STD of Normalized Friction Ratio & Classification Index U, Based upon log FR_1 & log q_{c1} .

successfully used in this issue so that satisfactory results can be obtained. This fact indicates that something more fundamental exists in the interpretation of current soil classification charts. Also, some geometrical meanings can be found for the soil classification index, U, and the soil in-situ state index, V. In the case of Olsen's chart, it can be proved mathematically that the U and V are two parameters, each of which controls a group of lotus functions of ellipses in log f_{s1} --- log q_{c1} space. Specifically, they decide the sizes and positions of those ellipses in that space. Soils satisfying same lotus functions will have same U or V values. The sought correlations between soil behaviors and soil types or other properties can be explained geometrically as the correlations between these ellipses and soil types or other properties.

1.7. Summary

It has been observed that there are two tendencies in most existing **CPT** or **PCPT** soil classification charts, where soil type changes in one direction and in-situ soil state changes in another. As the result of the research work presented in this chapter, all of these two-dimensional soil classification charts can be simplified into one-dimensional soil classification systems by a linear conformal mapping. Accordingly, Douglas et al's chart (1981), Olsen et al's chart (1988), and P. K. Robertson's charts (1990), as shown in Figure 1.1, 1.2, and 1.3, have been transformed into their corresponding U --- V planes as shown in Figure 1.6, 1.8, 1.10, 1.12. Based upon these transformations, the corresponding one-dimensional soil classification systems can be established by the soil classification index U and the new systems are at least equivalent to the original ones.

Furthermore, the generality of the simplifiability on current soil classification charts of cone reflects the fact that cone tip resistance q_c and friction ratio FR are not fundamental parameters even if they are normalized with respect to some factors. Therefore, reorganizing these parameters to obtain a primary indicator of soil types should be an essential step to develop an efficient soil classification system. This can be implemented by a conform mapping. One benefit of this reorganization is that the advantage of continuous and visual description of site condition in raw cone data is well preserved in the analysis results of site stratigraphy. Consequently, the efficiency of expressing the interpretation of cone data is greatly improved. Also the dimension of resulting new classification systems is reduced so that a possible statistical modelling of the certainty in these new systems will be much simpler.

In order to further illustrate the validation of the simplification approach suggested here, some **CPT** data from four sites have been used to check the two new classification systems of **CPT**, as shown from Figure 1.15 to Figure 1.24. The outcomes confirm that the classification results from Douglas' chart will be affected by the depths of soils but the results from Olsen's chart are free from this kind of influence, i.e., based upon the normalized soil behavior. These figures further indicate that only a statistical certainty can hopefully be achieved in these classification systems so that further research efforts should be put in the direction of a statistical evaluation on them.

One shortfall in this illustration procedure is that the implementation of the behavior unit concept is based upon a subjective process which is individual based and so is widely open to variation. In next chapter, this problem will be further discussed and an objective procedure of implementation based upon a statistic will be suggested.

CHAPTER 2

IDENTIFICATION OF SOIL STRATIGRAPHY

2.1. Introduction

The identification of soil layers and the classification of soil types are two basic aspects of determining soil in-situ stratigraphy when a conventional (crisp) soil classification system is used. The former is focused on the characteristics of spatial distribution of soil types. The latter then is interested in the characteristics of physical and engineering properties of soils. Approaches to these two aspects are closely related in general and this close relation becomes even stronger in the technology of interpreting cone testing data. The reason is that both of them rely on same set of cone testing data and so are intrinsically not independent from each other, as already shown in Chapter 1.

Theoretically, an identification of soil types over a sounding log will naturally result in a set of layers over the profile. However, the actual process of identification goes in an opposite way, i.e., identifies layers first and decides soil types second. This is true in the conventional technique of boring logs since, in this way, only a limited number of representative soil samples are needed to determine the corresponding soil types. Believed or not, it also has to be true in the interpretation of cone testing data. This chapter is dedicated to explain why the identification of soil layers should be done first in the process of interpreting cone data and how it can be accomplished reasonably. All the discussion will take **CPT** data as examples.

2.2. In-Situ Soil Behavior Unit

Soil classification systems based upon cone testing data have been considered as soil behavior type since their classification indices are relied on soil responses to cone penetration. These responses are cone tip resistance q_c and friction resistance f_s in **CPT** case, as already known in Figure 1.13. It has been noticed that only the values of q_c and f_s themselves are sometimes not enough to represent the normal behaviors of soils tested. The knowledge of their spatial distributions is also needed. Research works have already shown that different responses of q_c for same type and state of soils were observed in chamber tests simply because of soil layering with different thicknesses (**Schmertmann, 1978; Douglas et al, 1981; Robertson, 1984; Campanella et al, 1988**). Therefore, a concept of behavior unit is needed in order to describe the normal behaviors of soils subject to cone penetration in quantity. As already mentioned in Chapter 1, this behavior unit property should contain not only the information on the values of q_c and f_s, but also the information on the spatial distributions of them, i.e., the information of layering. Consequently, in the case of soil stratigraphy based upon cone testing data, determining soil layers still needs to be done first.

It is known from the layering experience obtained in Chapter 1 that as a fundamental step of a **CPT** soil engineering classification, a rational and objective layering procedure should be adopted to avoid possible human errors. A statistical method of layering, suggested by D. S. Wickremesinghe (**1989**), is a good candidate. In that method, a moving "window" with fixed width (W_d) is put over a given profile to analyze. The exposed portion of the data profile within the window is examined, with the central line, d_0, of the window being a potential boundary. The potential boundary line will divide the exposed data into two samples which are verified for distinctness. **Larger the difference between the two samples and less the variation within each of them, better is the chance for the corresponding central line of the window to be a real boundary line.** The measurement of this distinctness can be implemented by an Intraclass Correlation Coefficient (ICC), ρI. The definition of this coefficient and the discussion of it will be presented in the next section. The conclusion given here is that larger the distinctness of the two data samples and less the variation of the data within each of them, larger the value of ICC will be.

As the window moves along the given profile in steps equal to the sampling (or reading) spacing of the data, values of ρI are calculated for the potential boundary lines at the central line, d_0, of the moving window. All these values are then plotted against depth so that a set of most possible boundary lines of layers can be determined along the profile. Figure 2.1 presents an example of such a curve along the depth, where the corresponding U profile is also plotted. Naturally, a relatively high value of ρI for an ICC peak at a depth will indicate a presence of a layer boundary at that location.

Two decisions have to be made before this layering process can be used in CPT case. First, choose a candidate profile to work with. Since layering is based upon the information of soil types, not in-situ states, a profile of soil classification index, U,

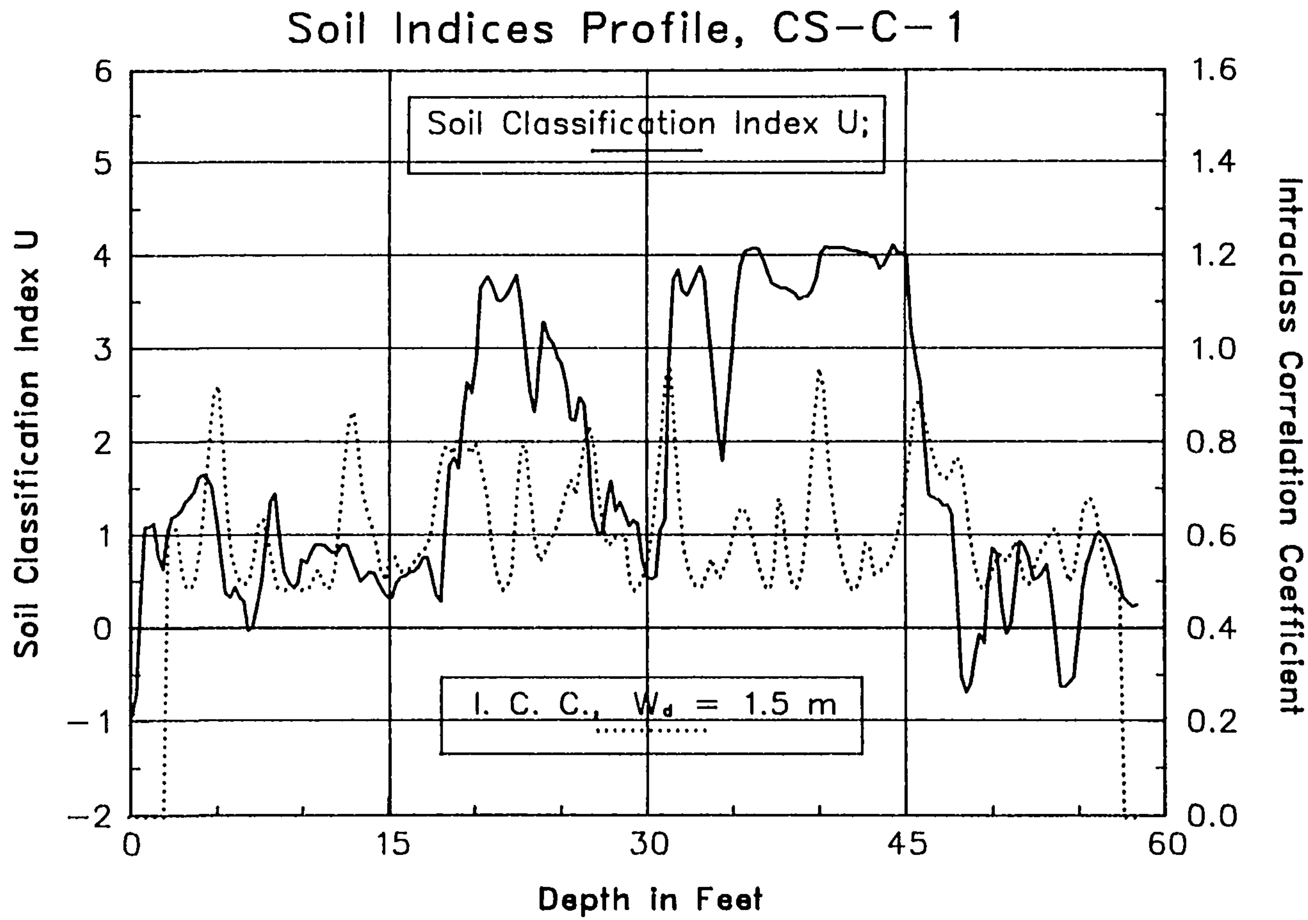


Figure 2.1 The U Profile of Figure 1.14 Overlapped with ICC Curve (W_d = 1.5 m).

should be used. In the terms of FR and log q_c, as an example, the classification index, U, is determined by merging Equation (1.3), (1.6), (1.7), and U = -u, as shown in Equation (2.1), (2.2), and (2.3).

$$x = 0.1539\,FR + 0.8870\log q_c - 3.35 \qquad (2.1)$$

$$y = -0.2957\,FR + 0.4617\log q_c - 0.37 \qquad (2.2)$$

$$u = -\frac{(a_1x - a_2y + b_1)(c_1x - c_2y + d_1)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} - \frac{(a_2x + a_1y + b_2)(c_2x + c_1 + d_2)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} \qquad (2.3)$$

here, $a_1 = -11.345$, $a_2 = -3.795$, $b_1 = 15.202$, $b_2 = 5.085$,

$c_1 = -0.269$, $c_2 = -0.759$, $d_1 = -2.960$, $d_2 = 2.477$.

Actually, Equation (2.1) and (2.2) are the same as Equation (1.6) and (1.7). Also, Equation (2.3) is the modification of Equation (1.3) by changing the sign. An instance of such a U profile with corresponding ρI is already exhibited in Figure 2.1. This U profile is from Figure 1.14.

Second, select the width of the moving window W_d. Two aspects, the physical and the mathematical, should be taken into consideration. Research has shown that "for the standard 10 cm^2 electric cone, the minimum stiff layer thickness to ensure full tip resistance is therefore between 36 cm to 72 cm. The tip may however, respond fully for soft layers considerably thinner than 36 cm in thickness" (**Campanella et al, 1988**).

Therefore, the W_d should conservatively take the value of 1.5 meter (m) so that the half of it will have the thickness of 0.75 m. The mathematical consideration is based upon the assumption that a wider window have a larger sample size. The effect of a sample size on ICC will be discussed in the next section, but as a conclusion here, the window width, W_d = 1.5 m, is a reasonable choice. The ρI profile shown previously in Figure 2.1 is calculated with W_d = 1.5 m. Several ρI curves with different W_d values for the same U profile will be shown in Figure 2.2 in the next section.

Now, an in-situ soil behavior unit in **CPT** can be defined as follows:

1). Use the **ICC** moving window method to layer a profile of soil classification index U;

2). Find the mean of U values for each layer.

These mean values are taken as the in-situ behavior units of corresponding soil layers and all so-determined soil layers therefore have a property of statistical homogeneity in this sense.

2.3. Intraclass Correlation Coefficient (ρI)

Given a moving window with the width of W_d over a candidate profile, there will be two samples, Ω_1 and Ω_2, on each side of the window central line, d_0. Let μ_1 and μ_2 be the means of the samples and σ_1^2 and σ_2^2 be the variances with n_1 and n_2 as the sample sizes, respectively. Here

$$\sigma_1^2 = \frac{1}{(n_1 - 1)} \sum_{i=1}^{n_1} (y_i - \mu_1)^2 \tag{2.4}$$

$$\sigma_2^2 = \frac{1}{(n_2 - 1)} \sum_{i=1}^{n_2} (z_i - \mu_2)^2 \tag{2.5}$$

y_i and z_i are the data values in each of the samples. If $n_1 = n_2 = n$ is assumed, a pooled combined variance, Y_w^2, can be defined as:

$$Y_w^2 = \frac{n}{2n - 1} \sigma_1^2 + \frac{n}{2n - 1} \sigma_2^2 \tag{2.6}$$

The between class variance Y_b^2 of these two samples with equal size, n, is the variance of the combined sample given by:

$$Y_b^2 = \frac{1}{2n - 1} \sum_{i=1}^{n} [(y_i - \mu)^2 + (z_i - \mu)^2] \tag{2.7}$$

Here, μ is the mean of all the y_i and z_i within the window. Now, the Intraclass Correlation Coefficient ρI can be defined by

$$\rho I = \frac{Y_b^2}{Y_b^2 + Y_w^2} \tag{2.8}$$

Some changes on Equation (2.8) are necessary before any discussion on the characteristics of the ICC. It can be proved that the between class variance Y_b^2 can also be written as (see Appendix B)

$$Y_b^2 = \frac{n-1}{2n-1}\left(\sigma_1^2 + \sigma_2^2\right) + \frac{n}{2\,(2n-1)}\left(\mu_1 - \mu_2\right)^2 \qquad (2.9)$$

Substituting Equation (2.6) and Equation (2.9) into Equation (2.8) and rearranging it, we get

$$\rho I = \frac{1}{1 + \dfrac{1}{\dfrac{n-1}{n} + \dfrac{(\mu_1 - \mu_2)^2}{2\,(\sigma_1^2 + \sigma_2^2)}}} \qquad (2.10)$$

It is clear from Equation (2.10) that the **ICC** is mainly controlled by the square of the difference of means between these two samples, $(\mu_1 - \mu_2)^2$, and the sum of variances of them, $\sigma_1^2 + \sigma_2^2$. Larger the square of the mean difference and less the sum of the variances, larger the **ICC** value (ρI) is. This property corresponds well with the criteria we used when layering a **CPT** profile. In a case where the square of the mean difference is very large and the sum of the variances is very small, i.e., the ratio m = $(\mu_1 - \mu_2)^2/(\sigma_1^2 + \sigma_2^2) >> 1$, the ρI will approach unity (1), which is the maximum value an **ICC** can take. On the other hand, if the ratio m is close to or even equal to zero, i.e., the difference between the two samples is not significant, the ρI value will approach to (n - 1)/(2n - 1), which is the minimum value an **ICC** can have.

The **ICC** is also influenced by the sample size n directly and indirectly. The direct influence is through the term (n-1)/n. Table 2.1 gives out some ρI = f(m,n) values for some sample size, n. The indirect influence is through the means and variances of the samples and has an important impact on the values of ρI. Wickremesinghe (**1989**) studied the sensitivity of window width (i.e. sample size n) and concluded that "the effect of a narrow window width (W_d) is the introduction of noise into the statistic under consideration. However, this would not affect the selection of layer boundary depths. In contrast, the choice of a wider window width could lead to the possibility of missing out possible layer boundaries."

Table 2.1. Intraclass Correlation Coefficient for Some n and m Values

Sample Size n	Intraclass Correlation Coefficient ρI			
	m = 0	m = 1	m = 5	m = 10
5	0.444	0.643	0.853	0.915
10	0.474	0.655	0.855	0.916
15	0.483	0.659	0.856	0.916
20	0.487	0.661	0.856	0.916
25	0.490	0.662	0.856	0.916
30	0.491	0.663	0.856	0.916
35	0.493	0.663	0.856	0.916
40	0.493	0.664	0.856	0.916

On the other hand, as the result of normal distribution restriction on samples, he recommended that "windows widths of less than 1.0 m not be selected" (**1989**). Combining this consideration with the physical constraint discussed before, the selection of W_d = 1.5 m is a reasonable choice. In such a case, the direct influence of sample size, n, on ρI is negligible as shown in Table 2.1 since n >= 15 if a reading is taken for every 5 cm or less. Figure 2.2 shows different ρI curves along depths based upon different W_d. These curves correspond to the U profile in Figure 1.14 and have verified the conclusion made by Wickremesinghe.

2.4. Summary

The identification of soil layers and the classification of soil types look like two separate tasks in the traditional site investigation due to the characteristics of boring test technique although they are related intrinsically and closely. In such a case, the identification of soil layers should be done first and the classification of soil types done next. Unfortunately, this patten has little change in the interpretation of cone testing data for a site investigation due to the characteristics of cone penetration. These characteristics have made the concept of soil behavior unit imperative in a cone soil classification process. Such a concept has in fact reflected the average responses of soils to cone penetration over a layer and can be implemented with the help of the moving window approach where, a statistic profile of Intraclass Correlation Coefficient (**ICC**), ρI, can be obtained. As a result, a set of soil layers over the U profile worked can be determined accordingly and each soil layer will have a property of statistical homogeneity. Therefore, soil behavior units for these soil layers can be decided, which will pave the way for establishing the correlation between soil behaviors and soil types.

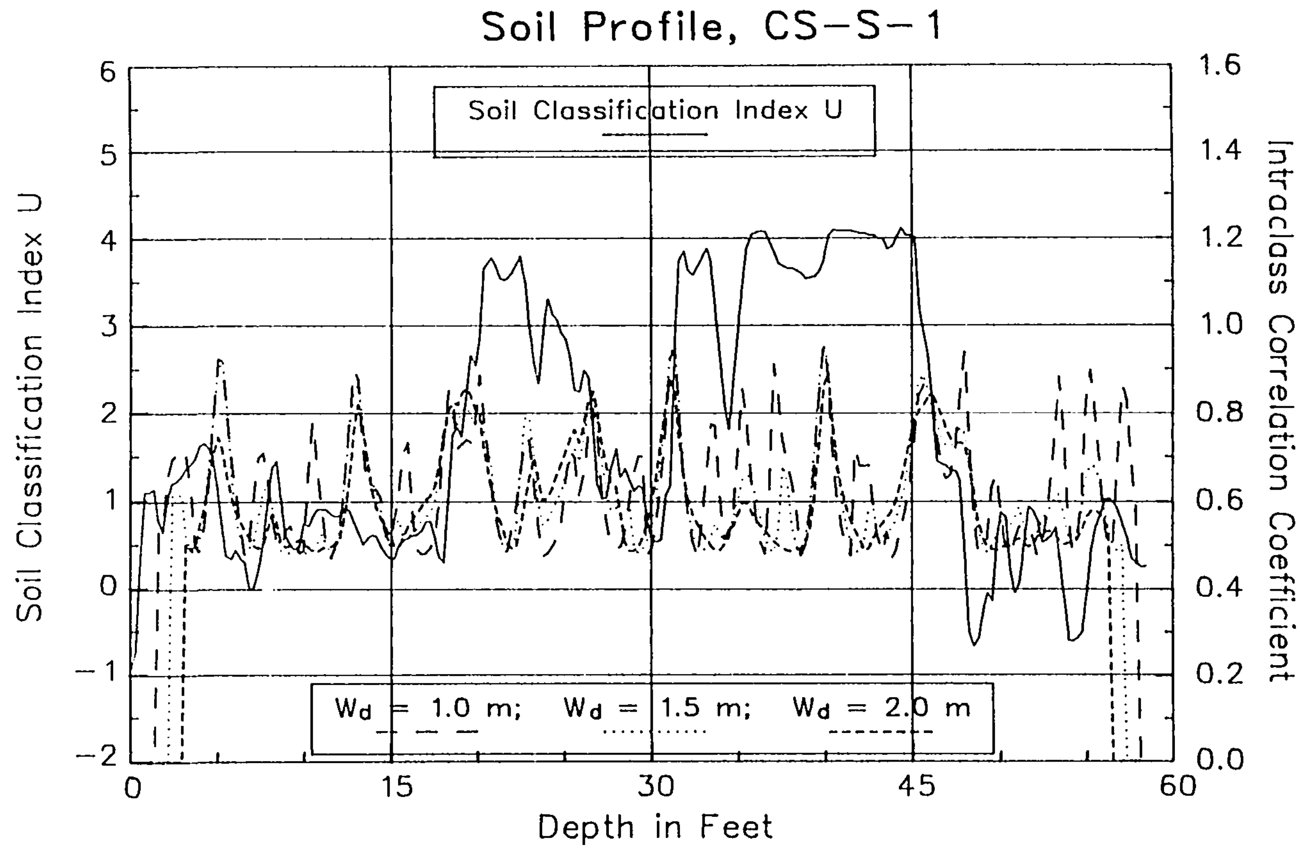


Figure 2.2 The U Profile of Figure 1.14 Overlapped with ICC Curves of Different Window Widths.

CHAPTER 3

UNCERTAINTY IN CPT SOIL CLASSIFICATION

3.1. Introduction

Soil classification, considered **the most confusing chapter** in soil mechanics by A. Casagrande (**1948**), is still confusing now and then when new technologies are employed to develop some new tools for soil engineering classification. Advances in science and technology have made cone penetration one of the increasingly popular in-situ tests due to its distinctive advantages. This test often provides better speed and economy, more detailed, repeatable and precise data, and data better suited to many ordinary soil engineering design problems (**Schmertmann, 1978**). The widespread use of it in engineering has requested a soil engineering classification system based on it. Consequently, several soil classification charts of this kind have been suggested. Either cone tip resistance q_c and friction ratio FR or their normalized forms have been used in these charts (**Schmertmann, 1978; Douglas & Olsen, 1981; Tumay 1985; Robertson et al, 1986; Olsen and Malone, 1988; Robertson, 1990**). Also a recent adaptation has included pore water pressure as one of classification chart parameters with various definitions (**Jones and Rust, 1982; Senneset and Janbu, 1984; Campanella and Robertson, 1988; Senneset et al, 1989; Robertson, 1990; Cheng-hou et al, 1990**).

Although the up-date classification charts have been improved substantially with the increasing knowledge of this testing technology in recent years, the overlaps of different soil types are still found in them (**Douglas & Olsen, 1981; Larson & Mitchell, 1986; Campanella & Robertson, 1988**). The partial reason for the unsuccessful effort to eliminate the overlaps is that they are inherent. The certainty in all these charts can only be statistical in nature. It is obvious that an adequate understanding and discussion of this overlapping phenomenon will help us to appropriately use current **CPT** soil classifications and to develop a better one.

3.2. Statistical Bias in Soil Classification of Cone Technology

It has been believed that some of the most comprehensive recent work on soil classification using **CPT** data was presented by Douglas and Olsen (**Robertson, 1990**). According to their article published in 1981, **USCS** has been used as classification criteria to determine soil types in the process of developing a **CPT** Classification Chart (**CPTCC**). This approach has also been adopted by other investigators. Thus, a **CPTCC** is based upon a correlation between the behaviors of soils subject to cone penetration and the soil composition, and has inherited the same classification criteria used by **USCS**. This approach has the advantage of keeping the continuity and integrity of the whole soil classification system so that the accumulated experience related to **USCS** is still available to the users of the **CPTCC**. However, some uncertainties will be inherited or introduced, too.

USCS was originally developed by Professor A. Casagrande during World War II for use in airfield construction, on the basis of analyzing the existing experience and soil classification systems to that date. "It was modified in 1952 by Professor Casagrande, the U.S. Bureau of Reclamation, and the U.S. Army Corps of Engineers to make the system also applicable to dams, foundations, and other construction. The basis for the USCS is that coarse-grained soils can be classified according to their grain size distributions, whereas the engineering behavior of fine-grained soils is primarily related to their plasticity. In other words, soils in which 'fines' (silts and clays) do not affect the engineering performance are classified according to their grain size characteristics, and soils in which fines do control the engineering behavior are classified according to their plasticity characteristics" (**Holtz, Kovacs, 1981**).

It was emphasized by Casagrande (**1948**) that this system was not merely to classify the raw material of which soils were composed without regard to engineering properties of the material in its undisturbed state. Compositional factors and engineering properties of the material "must be considered as an integral part of the complete classification" (**Casagrande, 1948**). However, because of the complicated relations between them and also the comprehensive nature of soil classification which is based upon the experience of a long period of time, "an exact point of division is unimportant" (**Casagrande, 1948**) for some boundaries of the system. "It is intended that the classification of soils in accordance with this system have some degree of elasticity, and that the system not be followed blindly nor regarded as completely rigid" (**Waterways Experiment Station, Vicksburg, Mississippi, 1951**). Therefore, USCS "is for qualitative application only" (**ASTM Designation: D 2487-85**), and "when quantitative information is required for detailed designs of important structures, this test method must be supplemented by laboratory tests or other quantitative data to determine performance characteristics under expected field conditions" (**ASTM Designation: D 2487-85**).

Two points can be obtained from this brief review of the **USCS** history. First, **USCS** is a well established soil engineering classification but the correct usage of it in a site investigation can not be guaranteed by a mechanical process of soil classification. Some expertise must be utilized in both the soil classification itself and the way to determine soil layers and to obtain soil samples. This expertise is individually based and open to a variation so that it is possible to have some disparity in soil classification results if they are from different sources.

Second, it is known that in a compositional soil classification, such as the **USCS**, soils are categorized according to their compositional characteristics and probable engineering behaviors. However, when the type of a soil sample is known, its engineering behavior can be predicted only in quality, not in a precise quantity. Any compositional soil classification will not eliminate the need for detailed soil investigations to quantitatively predict engineering properties of soils. This reflects the fact that the correlation between the soil composition and soil behaviors is only statistical in nature. Most correlations found between soil compositional factors and soil behavior properties were based upon average values and average behaviors. As a specific case, it is not surprising that the correlation between responses of soils subject to cone penetration and soil types has this pattern, too. Therefore, the overlapping in current **CPT** soil classification charts is unavoidable and the anticipation of soil types based on **CPT** will definitely have some uncertainty.

Apart from the uncertainty caused by using compositional criteria, systematic researches on cone penetration technology itself have already shown that the data from cone tests will be affected by many factors (**Douglas & Olsen, 1981; Campanella & Robertson, 1988**). Among them are equipment and procedure, soil composition, and environmental factors. These three primary components can simultaneously interact during a sounding (**Douglas & Olsen, 1981**). With the accumulation of knowledge on cone penetration technology, the importance of standardization on equipment and procedure of this test has been realized. Therefore, a standard international reference test procedure for cone penetration has been established (**E. E. De Beer et al, 1988**). It is now possible to maintain a rigorous standardization of cone penetration, which will make **CPT** data from different sources comparable. Nevertheless, the standardization can not eliminate the probabilistic variation of cone testing data due to the complex nature of the technique for testing.

Furthermore, on the condition that purely equipment- and procedure-related influences can be eliminated from concern by standardization, soil composition and environment factors and their interaction with a penetrometer are still in question. Their contribution to the measurements of cone penetration can only be qualitatively explained as follows.

"Within coarse-grained materials an increased grain size, or well graded distribution, or angularity of particles all increase end bearing at low friction ratios. Fine-grained materials show lower end bearing at higher friction ratios with increasing plasticity. An increased end bearing at a given depth will result from an increase in strength-effective stress ratio. ... Increasing density causes an increase in end bearing, as does an increase in confining pressure, overconsolidation ratio, or coefficient of lateral pressure. The effect of these factors on sleeve resistance is not as well known" (**Douglas & Olsen, 1981**). "The measured pore pressures are influenced by factors, such as, stress history, sensitivity and stiffness to strength ratio (G/S_u)" (**Campanella & Robertson, 1988**).

"Since theories to truly model the penetration process for all soils are still in the process of development, much disagreement still exists over the distribution of stresses around a penetrating cone" (**Campanella & Robertson, 1988**). Consequently the lack of theoretical analysis on results of cone penetration still exists. Also still, no clear theoretical explanation is available now for the relationship between soil types and soil behaviors exhibited during cone penetration. As a result, the interpretation of correlations between them is still fundamentally empirical. All these will bring some uncertainty into current **CPT** soil classification charts. Naturally, it has been stressed that "these charts are global in nature and should be used only as a guide to define soil behavior type" (**Campanella & Robertson, 1988**).

As a summary, it can be concluded that soil engineering classification based upon cone penetration technology has some uncertainty. Several reasons can be accounted for:

- The random nature of soils;
- The correct usage of **USCS** in a site investigation is based upon expertise which varies individually. If boring data are from different sources, it will be possible for them to have some disparity on soil classification results;
- Relations between compositional parameters and behavior measurements of soils are statistical or variational in nature;
- The test procedures of cone penetration technology are much more complicated than conventional laboratory classification index tests;
- The data from cone penetration are influenced by test equipments and procedures, soil compositions, and environments;
- The interpretation of the data from cone penetration is mainly based on an empirical approach.

Moreover, it is the advantages of cone penetration technology and "the outstanding disadvantages of not obtaining a soil sample for visual and lab inspection" (**Schmertmann, 1978**) that have urged us to generate a pertinent procedure to predict soil types. In other words, the choice of cone testing data as soil classification parameters is not fully because they are as good indicators on soil types as grain size and plasticity characteristics. The researching efforts on the selection of soil classification indices based upon cone testing data itself is the proof that this is not a solved problem.

It is apparent that the disadvantages of cone penetration technology has already prevented us to directly use **USCS**. It also prevents us from developing a soil classification system which, when used, will present some kind of uncertainty (i.e., some degree of elasticity) but will be supplemented by visual, lab inspection and personal experience. Therefore, as a step to compensate for the weaknesses of this technology, a statistical modelling of the uncertainty in **CPT** soil classifications will help engineers understand and properly use **CPT** soil classification data. This topic will be fully expanded in Chapter 4. As a preparation here, the next section in this chapter will discuss a preliminary data reduction on some in-situ testing data.

3.3. Characteristics of Soil In-situ Behavior Units

CPT data from eight (8) testing sites, as shown in Table 3.1, have been used to illustrate the uncertainty in the correlation between soil types and soil in-situ behavior units of **CPT** and to provide a data basis for a suggested statistical modeling on the uncertainty. For data from a cone with a cross-sectional area of 15 cm^2, Equation (1.16) and (1.17) are used in order to get the equivalent data comparable with the ones from a cone with a cross-section area of 10 cm^2. In order to make the **CPT** data from each site have an equal weight in a possible statistical analysis, only three (3) **CPT** sounding logs are chosen from each testing site. These selected **CPT** data are pre-treated in the ways

Table 3.1. Brief Information about **CPT** Data Used

Site Name	Number of CPT Sounding	Number of Boring Log	Cone Type	Data Source
San Diego	9	8	I	Ref. 1
Salinas	15	8	I	Ref. 1
Moss Landing	8	8	I	Ref. 1
San Jose: Coyote South	11	4	I	Ref. 1
San Jose: Coyote North	9	4	I	Ref. 1
Alabama, Theodore S.	9	8	II	Ref. 2
Alabama, Theodore N.	6	12	II	Ref. 2
Louisiana, South	18	18	II	Ref. 3

Note:

Type I: A tension type cone from Fugro with a cylindrical friction sleeve of 150 cm^2 surface area, capped with a 60-degree apex angle conical tip of 10 cm^2 projected surface area.

Type II: A 43.7 mm nominal diameter Fugro cone penetrometer (cross-sectional area of 15 cm^2), with a friction sleeve area of 200 cm^2 and a cone apex angle of 60 degree, being a substraction type.

Ref. 1: USGS Open-File Report No. 81-284, 1980. "Evaluation of the Cone Penetrometer for Liquefaction Hazard Assessment", Prepared by Fugro, Inc..

Ref. 2: Data from Southern Earth Science Inc. and the Department of Civil Engineering, Louisiana State University.

Ref. 3: Data from Louisiana Transportation Research Center.

suggested previously in Chapter 1 and 2 to obtain soil behavior units and the corresponding soil types are determined according to the corresponding boring log profiles.

The suggested preliminary data reduction consist of two steps. The first step is to reorganize parameters FR and log q_c to obtain the corresponding profiles of the soil classification index U by using Equation (2.1), (2.2), and (2.3). The normalized soil classification parameters after Olsen (1988) are not used here due to their iterative nature. It is difficult to implement the iterative process to determine those parameters in a statistical analysis.

The second step is to use the method of ICC moving window suggested in Chapter 2 to determine soil in-situ behavior units over those profiles of soil classification index U. The empirical procedure to layer a profile can be described as follows. First, the corresponding ICC profile of a U profile is determined and then the positions of all the peak values of ρI in this ICC profile are checked out. If a peak value of ρI in the ICC profile is equal to or larger than 0.7 ($\rho I_{peak} >= 0.7$), a boundary line is supposed to exist at that peak location. Otherwise, that location will be skipped. This process of determining boundary lines proceeds from the top down to the bottom of the ICC profile, and can be called as the primary layering.

Practice has already shown that only the primary layering will not give out satisfactory results of layering. The reason is that the primary layering sometimes will lead some profiles to have too many layers and others too few. Such a situation later will cause bias in a statistical analysis just because of the big differences of layer thicknesses. If a soil layer is too thin, it will not present the normal behavior of that soil type, but if a soil layer is too thick, a part of information contained in this layer will be wasted since it is treated only as one sample value in a statistical analysis.

Therefore, an empirical addition over the primary layering is taken as follows. When a boundary line is selected, the thickness of the layer made by this boundary line is checked. If the thickness is less than 1.5 feet (approximate 0.45 m), this boundary line will be discarded. On the other hand, if the thickness is larger than 7.5 feet (2.3 m), that layer should be further divided by seeking other peak **ICC** values (< 0.7) within that layer, which is called secondary layering. In this way, some new sub-boundary lines will be inserted into that layer at corresponding positions. This process continues until all sublayers are larger than 1.5 feet (0.45 m) and less than 7.5 feet (2.3 m). No **ICC** value constraint is imposed on this later case (secondary layering).

Based upon the layering procedures discussed above, the selected **CPT** data from the eight (8) sites are manipulated accordingly. The resulting relative frequencies of layer thicknesses with or without thickness constraint are shown in Table 3.2 for the purpose of comparison. It can be seen from this table that the relative frequency of layer thicknesses with thickness constraint is closer to a normal distribution than the one without the constraint. It is expected that the bias caused by the big difference of soil layer thicknesses will subsequently be reduced when the thickness constraint is applied.

Table 3.2. Relative Frequencies of Layer Thickness

Layer Thickness (ft)	Relative Frequency	
	without thickness constraint	with constraint
at or below 1.5	0.036	0.010
1.5 --- 2.5	0.218	0.165
2.5 --- 3.5	0.182	0.220
3.5 --- 4.5	0.160	0.192
4.5 --- 5.5	0.105	0.186
5.5 --- 6.5	0.051	0.150
6.5 --- 7.5	0.054	0.072
above 7.5	0.194	0.005

After the twenty four (24) **CPT** soundings from the eight (8) sites have been layered, the mean values of soil classification index U for these layers, known as soil in-situ behavior unit, are calculated. Their corresponding soil types are then determined according to the corresponding data of boring logs. Therefore, the basic results of the preliminary data reduction is a set of soil behavior units with corresponding soil type information. Based upon these results, a discussion on the characteristics of soil in-situ behavior units can be performed. All the results are plotted from Figure 3.1 to Figure

3.4. The data of four soil types are presented in this figures. They are sand (SP), silty sand (SM), lean clay (CL), and fat clay (CH).

The most significant characteristic of the **CPT** in-situ soil behavior units is that different soil types can have same in-situ behavior (i.e., same U values). Consequently, some overlaps among adjacent soil types will occur in a resulting **CPT** soil classification. It is this kind of overlaps that is the uncertainty discussed previously in a **CPT** soil classification.

Some factors can be excluded from inducing these overlaps. Figure 3.1 shows the scatter distribution of the soil behavior units along the depth for each type of soil. It can be seen that the influence of depth on soil behavior units (also represented by U) can be neglected. One possible reason might be that for each type of soil, the variations of responses to cone penetration can be divided into two kinds: intra-site and inter-site. The data currently available have already shown that the intra-site variation of cone responses is affected by the depths of soil layer positions. However, it also appears true that the inter-site variation of cone responses takes a dominant role in the total variation of cone responses. Therefore, when the analysis results from eight (8) test sites are put together, the influence of depth can not be seen anymore.

Another possible reason might be that the influence of depth (or confining stress) on soil responses to cone penetration has been totally included in (or transformed into) the soil

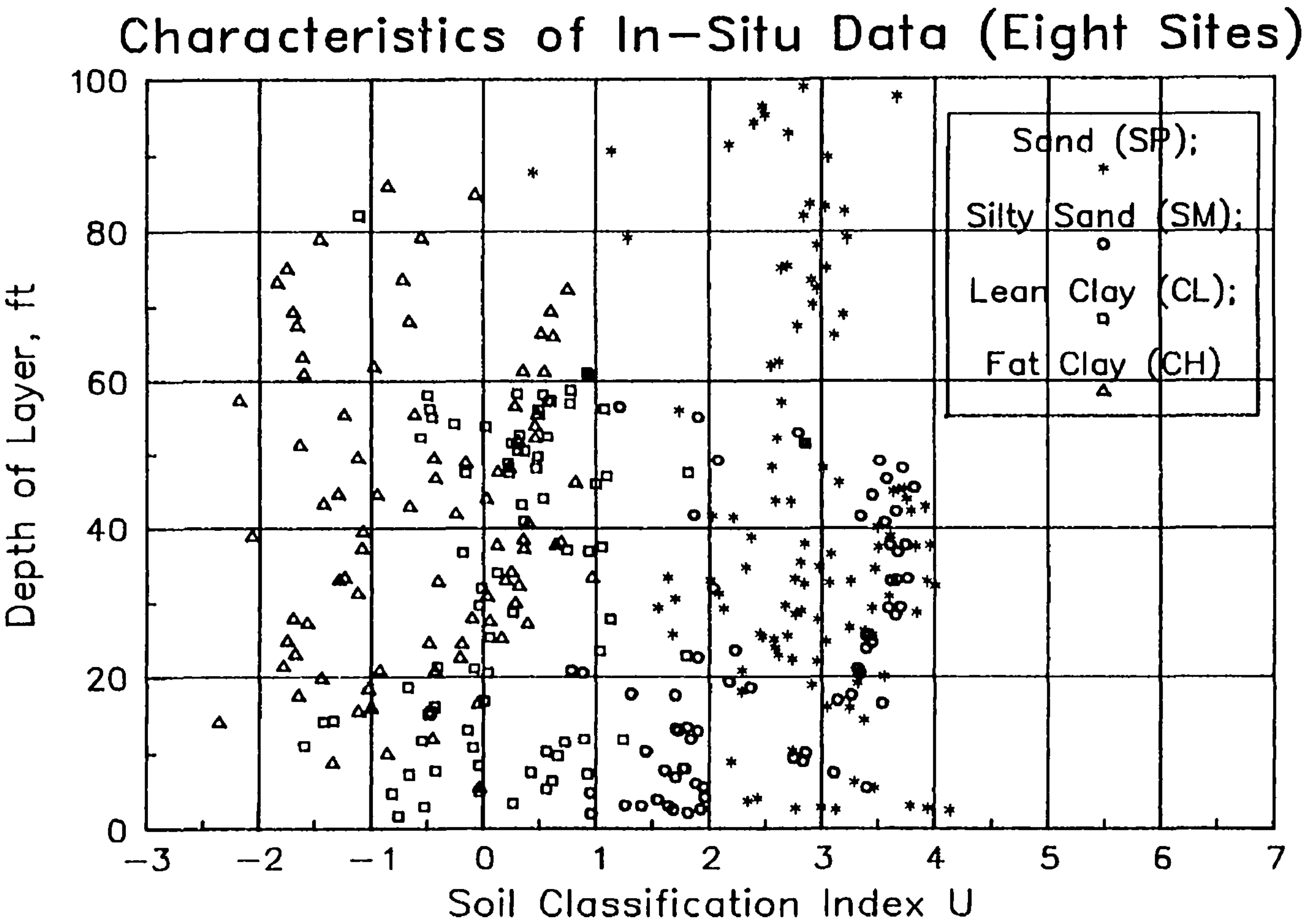


Figure 3.1 Scatter Distribution between Soil Behavior Unit U and Depth of Soil Layer.

in-situ state index, V. As a result, the soil classification index, U, is independent of it. If this is true, a **CPT** soil classification based upon the soil classification index, U, should be applicable to **CPT** data from different sources (in-situ testing, laboratory testing). No matter which reason it might be, a normalization with respect to the depth of soil layers is not needed at all in this context. Therefore, it can be concluded that different confining stresses are not the reason that causes those overlaps in Figure 3.1.

Figure 3.2 shows the scatter distribution of layer thicknesses for each type of soils. This figure supplements the layering results in Table 2.1 and indicates that layer thicknesses will not cause those overlaps, too. Figure 3.3 and Figure 3.4 present the scatter distribution of relative STDs of tip resistance and friction ratio corresponding to soil behavior units, separately. The relative STD of tip resistance is defined as (STD of q_c)/(average of q_c) and the relative STD of friction ratio as (STD of FR)/(average of FR) for each soil layer. These two figures have visualized the meaning of a statistical homogeneity.

Figure 3.5 and Figure 3.6 are organized in order to make the general characteristics of **CPT** soil behaviors more understandable. They exhibit the relative frequencies of soil classification index, U, and soil in-situ state index, V, for each type of soils. The sample size of each soil type is also given in them. It should be emphasized here that the U and V values are layer based. It can be seen that the relative frequency distribution of soil classification index, U, is dependent on soil types, but this

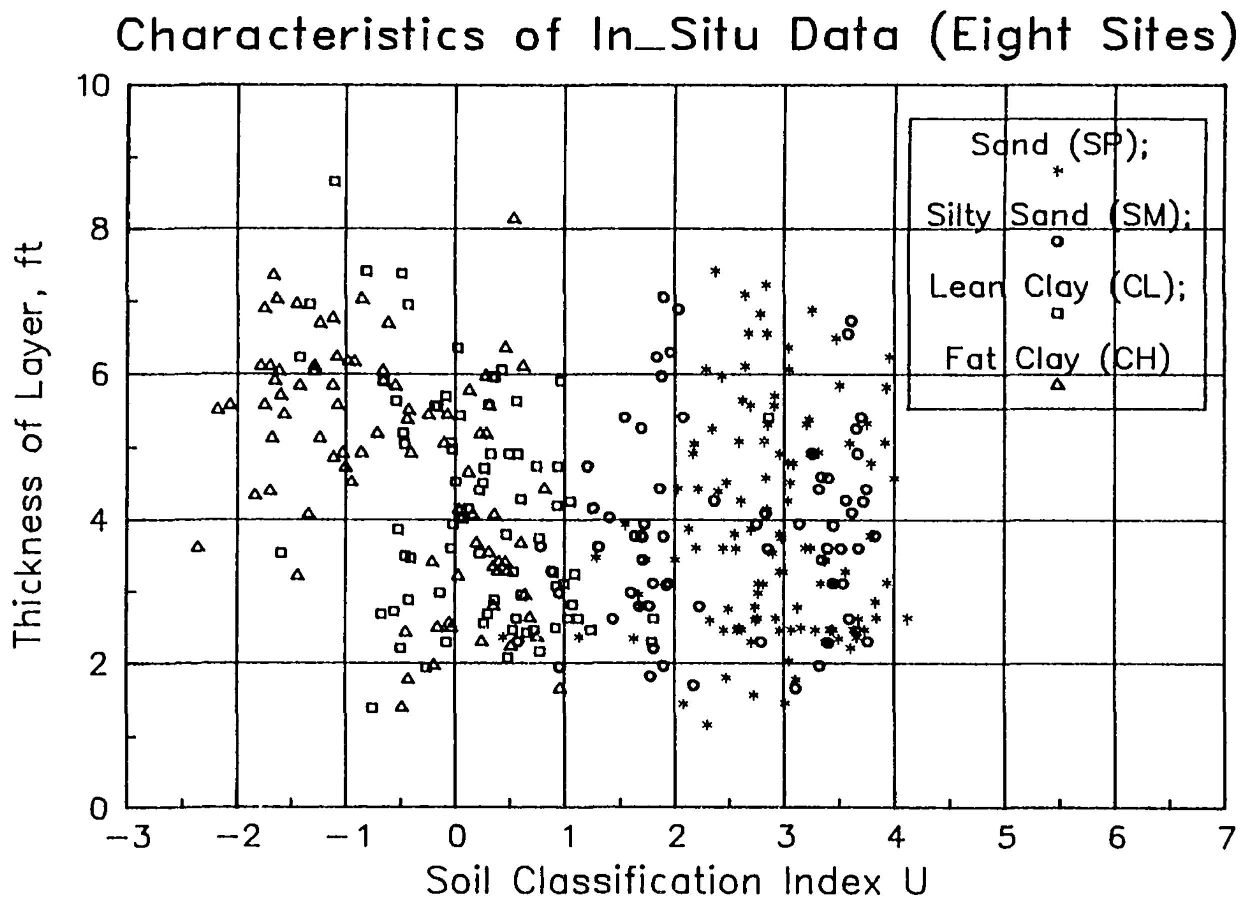


Figure 3.2 Scatter Distribution between Soil Behavior Unit U and Thickness of Soil Layer.

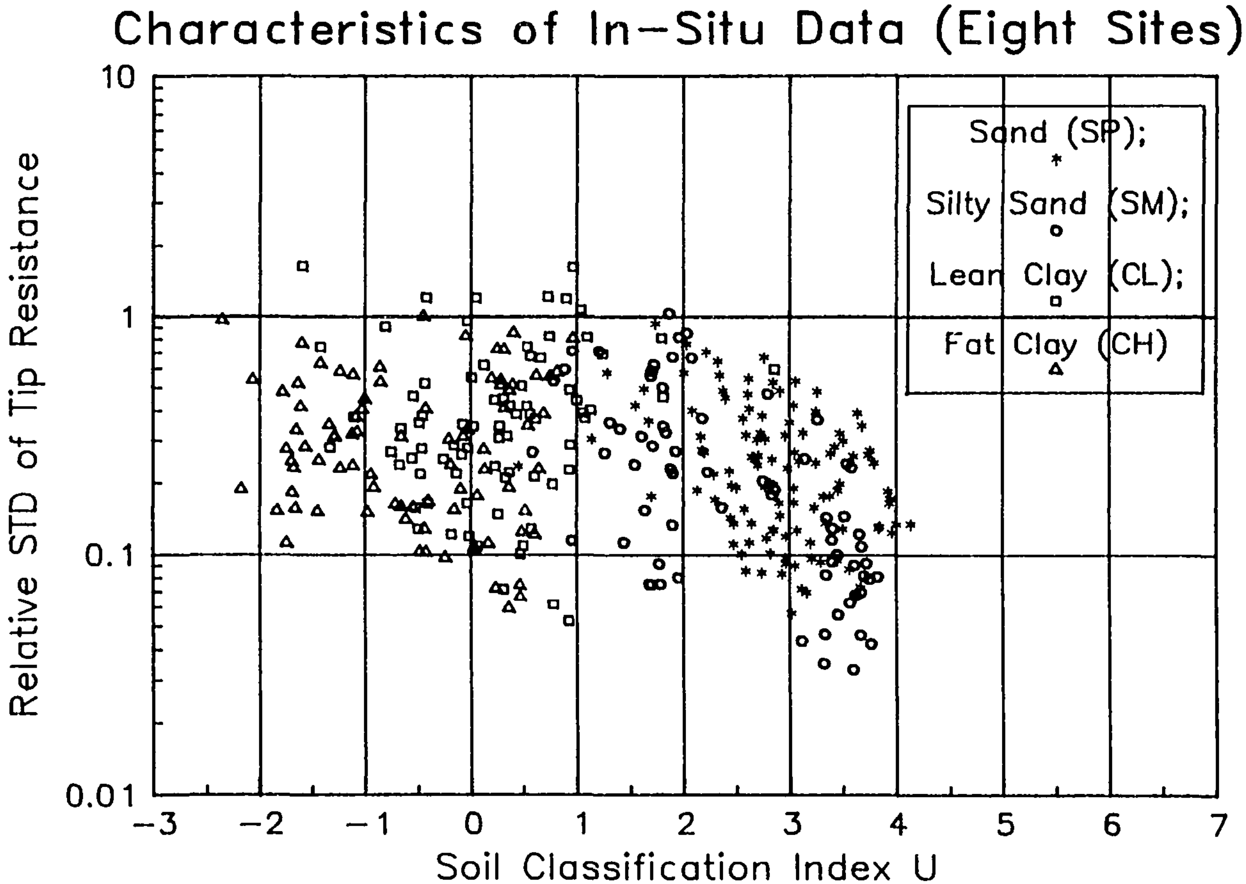


Figure 3.3 Scatter Distribution between Soil Behavior Unit U and Relative STD of Tip Resistance.

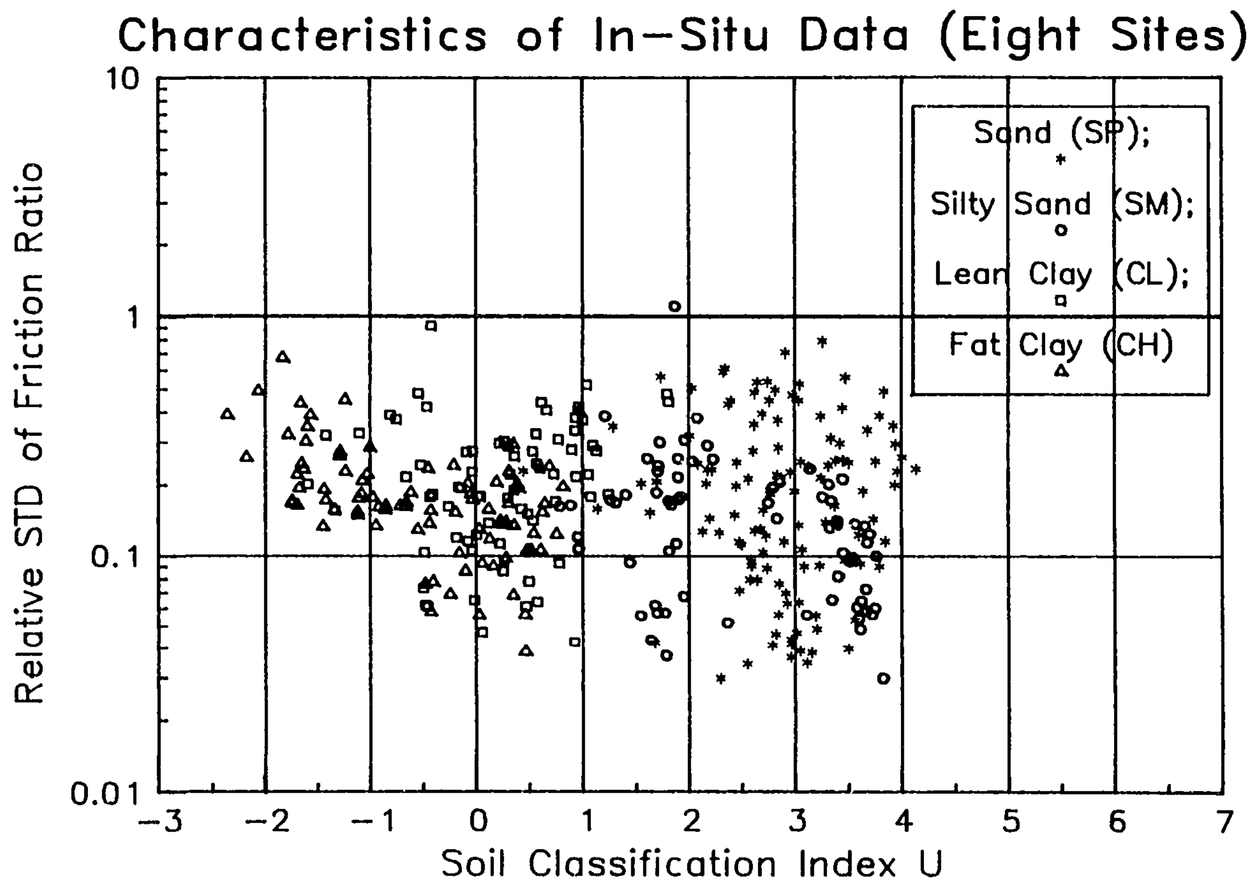


Figure 3.4 Scatter Distribution between Soil Behavior Unit U and Relative STD of Friction Ratio.

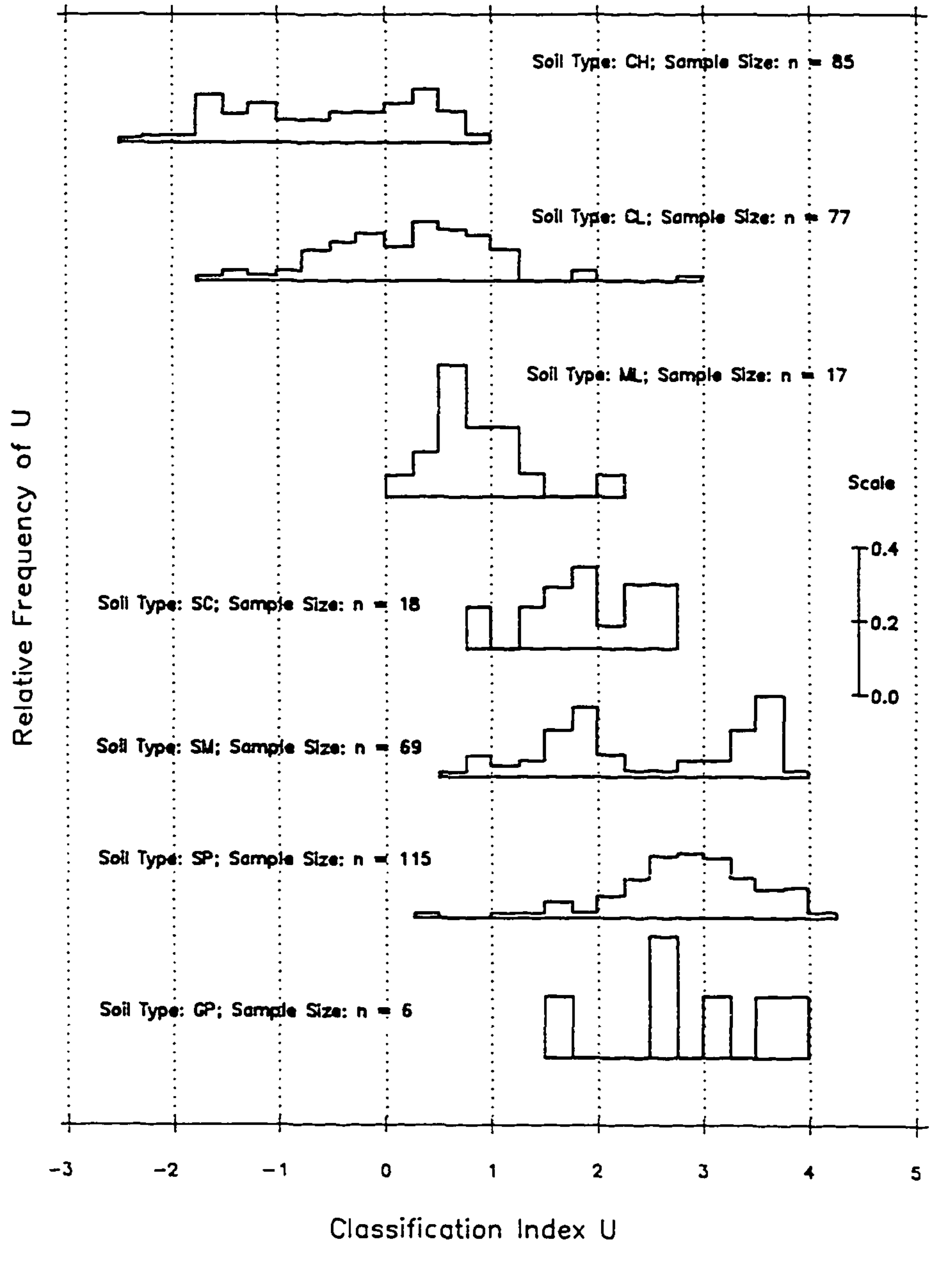


Figure 3.5 Relative Frequency of Soil Behavior Unit U for Different Soil Types.

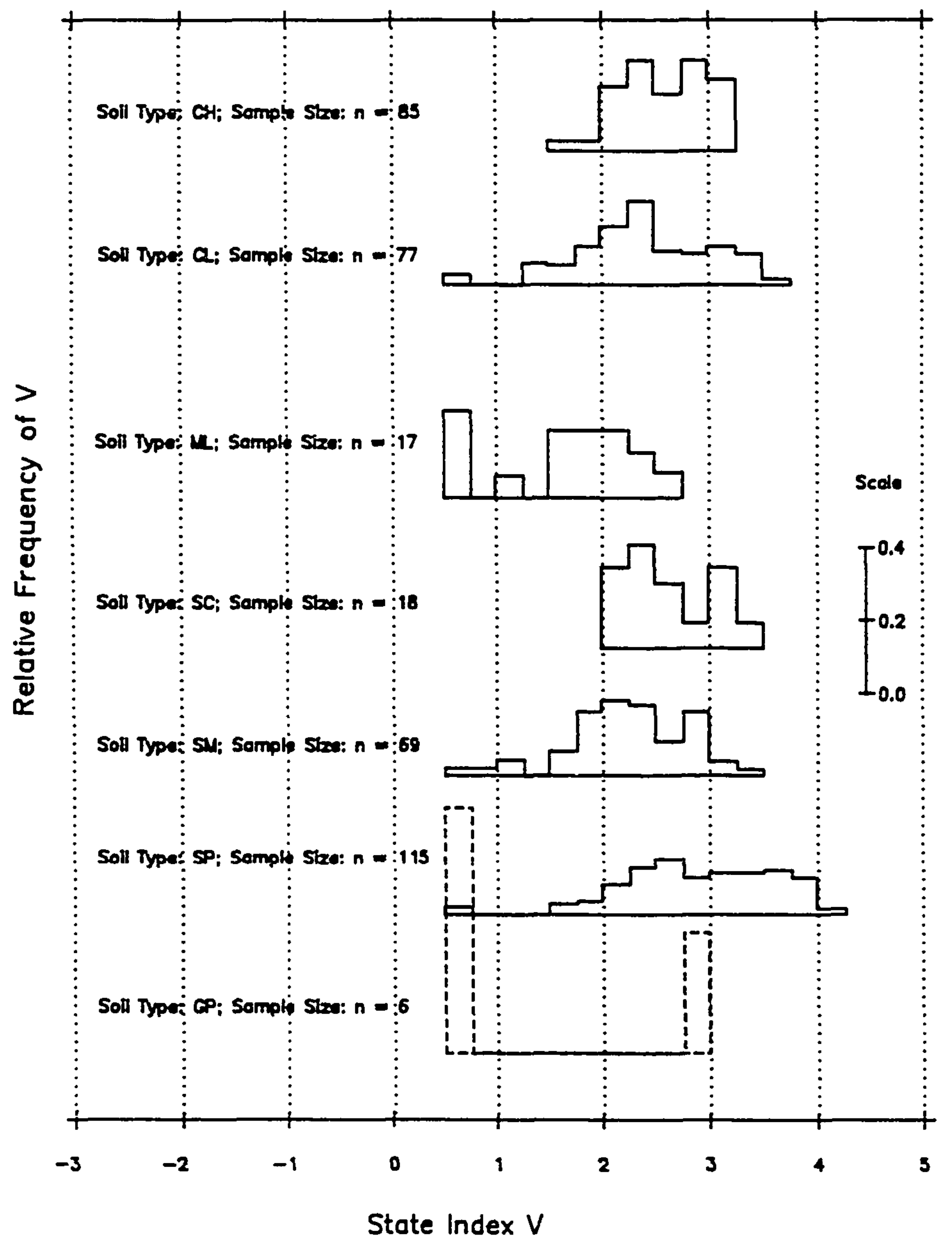


Figure 3.6 Relative Frequency of Soil In—Situ State V for Different Soil Types.

dependence is not defined very well due to the fact that overlaps exist. In other words, the diagram in Figure 3.5 actually shows us two facts. Different soil types will in general have different behaviors, but they sometimes can also have similar ones. Current **CPT** soil classification charts are trying to make use of the first fact, but the second fact causes the overlap or uncertainty problem. On the other hand, the relative frequency distribution of soil in-situ state index, V, is clearly independent of soil types, which is what we expected.

Now, an empirical **CPT** soil classification reference can be obtained if it is assumed that the relative frequency distributions of U given in Figure 3.5 are independent of the soil sample size, n. The relative frequency distributions of U are rearranged as shown in Figure 3.7 and Figure 3.8. An outstanding characteristic of these two figures is that the boundaries between different soil groups are actually banded shapes with different widths. In other words, there are overlaps between adjacent soil groups with different degrees. As an approximation, the middle points of the boundary bands in Figure 3.7 and 3.8 could be taken as the boundaries. In Figure 3.7, the boundary will be U = 1.25 for soil type group I (ML, CL, CH) and soil type group II (SC, SM, SP). If other information, such as the width of the boundary band and the degree of the overlap, is considered, it can be said that sandy and clayey soils will be identified quite well by **CPT** data.

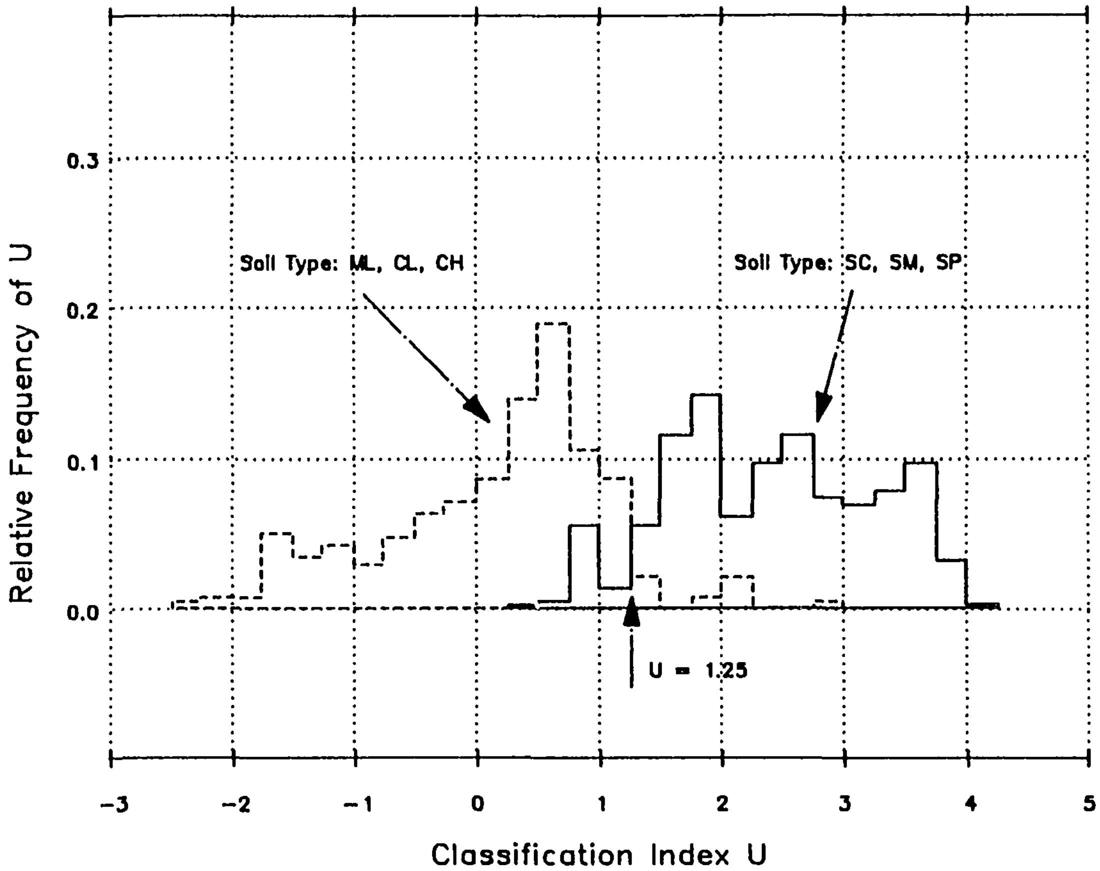


Figure 3.7 Empirical CPT Soil Classification for Two General Groups; Group I: SC, SM, and SP; Group II: ML, CL, and CH.

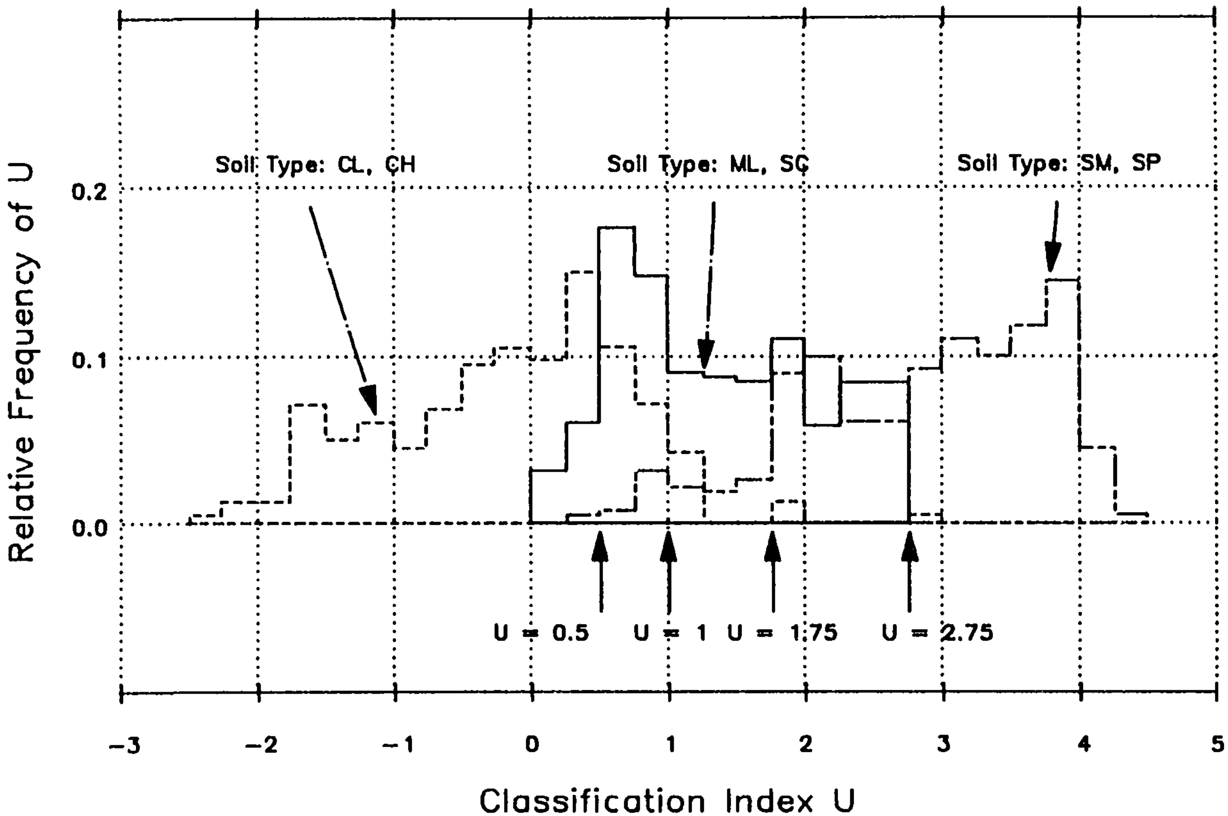


Figure 3.8 Empirical CPT Soil Classification for Three General Groups; Group I: SM and SP; Group II: ML and SC; Group III: CL and CH.

However, if the soil types are grouped in the way shown in Figure 3.8, the boundary values can not be defined with confidence. It might take U = 0.5 as the boundary between group I (CL, CH) and group II (ML, SC), and U = 2.25 as the boundary between group II and group III (SM, SP). An important fact, that should be kept in mind, is that if these boundaries are used as criteria to classify soils, the results of the prediction will not be one hundred percent (100 %) correct due to these overlap bands between different soil groups. Obviously, this conclusion is also true for the cases where individual soil types are predicted. All the results obtained so far have indicated the overlaps among different types of soils. However, it is not enough to only know the existence of the uncertainty. The knowledge of the probabilities with which the overlaps will occur is more important. How to handle this problem will be the subject fully discussed in Chapter 4.

3.4. Summary

Soil classification systems of cone technology still have an uncertainty of randomness although they have been updated substantially by an increasing knowledge. The following reasons can be found for this phenomenon:

- The nature of soils is random;
- The correct usage of **USCS** in a site investigation is based upon expertise which varies individually. If boring data from different sources are used, it will be possible for them to have some disparity on soil classification results;

- Relations between compositional parameters and behavior measurements of soils are statistical or variational in nature;
- The test procedures of cone penetration technology are much more complicated than conventional laboratory classification index tests;
- Data from cone penetration are influenced by test equipments and procedures, soil compositions, and environments;
- The interpretation of data from cone penetration is mainly based on an empirical approach.

Due to all these causes, it seems that a possible solution to the problem of uncertainty can only be worked out by a proper modeling of the uncertainty. Therefore, a preliminary data reduction is performed on the raw adopted **CPT** data from eight (8) testing sites according to the procedures suggested in Chapter 1 and 2 and a set of representative **CPT** soil behavior units with corresponding information of soil types is obtained. Based upon this intermediate result, an empirical correlation between soil types and soil **CPT** behavior units is established and a soil classification reference is suggested.

Two facts can be concluded from the analysis results obtained so far. In general, different types of soils have different behaviors, but sometimes they can also have similar ones, which again conform the uncertainty in the correlation between soil types and soil in-situ behaviors. Current **CPT** soil classification charts are trying to use the first fact, but the second one will cause the overlaps (uncertainty) in the charts. These overlaps are generally independent of depths and thicknesses of soil layers so that no corrections based upon them are necessary. This is one advantage of using the concept of soil behavior unit defined in this research. Furthermore, it is obviously not enough to only know the existence of the uncertainty. The knowledge of the probabilities with which the overlaps will occur is more important. How to determine these probabilities will be the subject fully discussed in Chapter 4.

As an extra point, soil classification criteria used to date are only good for compositional soil classification indices. Sticking to these classification criteria while developing a new soil engineering classification of cone technology will introduce or cause the uncertainty of randomness in the results of soil type predictions. The obstacle of this uncertainty can be avoided if a set of new independent classification criteria is adopted in a soil classification suggested. When such a new system is developed, the knowledge of the relations between the new and old soil classifications will be vital for users to have an access to both of them simultaneously. Obviously, it will be more desirable if the new soil classification includes such information.

CHAPTER 4

STATISTICAL PREDICTIONS OF SOIL TYPES FROM CPT

4.1. Introduction

Evidences have already shown that soil type prediction based upon **CPT** can only give a statistical certainty. This certainty may be described by some statistical models and will take a format of soil type predictions along with corresponding probabilistic statements. This chapter will discuss these statistical models according to the basic principles of probabilistic theory and propose some critical values for them. These models will provide tools to determine the probabilities with which different types of soils fall in the overlap ranges of a **CPT** soil classification discussed previously. Therefore, the discussion in this chapter will help users to correctly interpret and use the soil classification information from cone test data.

4.2. Normal Distributions of U for Different Soil Types

Statistical predictions of soil types require a knowledge of probability distributions of soil classification index, U, for each type of soils. Here, the index U represents in-situ soil behavior units defined previously in this research and is treated as a random variable. A simple and intuitive assumption on the probability densities of this random variable U for each type of soils is that each of these density functions has a normal distribution form determined by two parameters: a mean μ and a standard deviation σ.

This assumption is then checked by the in-situ soil behavior units of each soil type obtained previously in Chapter 3, as shown from Figure 4.1 to Figure 4.14.

Figure 4.1, 4.3, 4.5, 4.7, 4.9, 4.11, and 4.13 are normal probability plots, one for each type of soils. Each plot consists of an arithmetic horizontal axis and a vertical axis scaled so that the cumulative distribution functions of a normal distribution will plot as a straight line. All the figures can help us graphically assess whether the in-situ soil behavior units observed for each type of soils could reasonably have come from a normal distribution. Also, a least squares regression line is given out in each plot for comparison.

Figure 4.2, 4.4, 4.6, 4.8, 4.10, 4.12, and 4.14 display the frequency histograms (i.e., the observed relative frequency) of different soil types with the correspondingly fitted distributions (i.e., the expected relative frequency) superimposed on them. Table 4.1 is the results of distribution fitting tests based upon a Chi-square test and a Kolmogorov-Smirnov one-sample test. A blank entry in the table means the corresponding sample size is too small to perform the test.

The results of Chi-square test indicate that the soil samples from soil type SP and CL do follow the normal distributions with a quite high confidence, represented by α. However, the confidence for this claim is very low (α = 0.0045) in this scale for the sample from soil type CH and the assumption has to be rejected for soil type SM. The

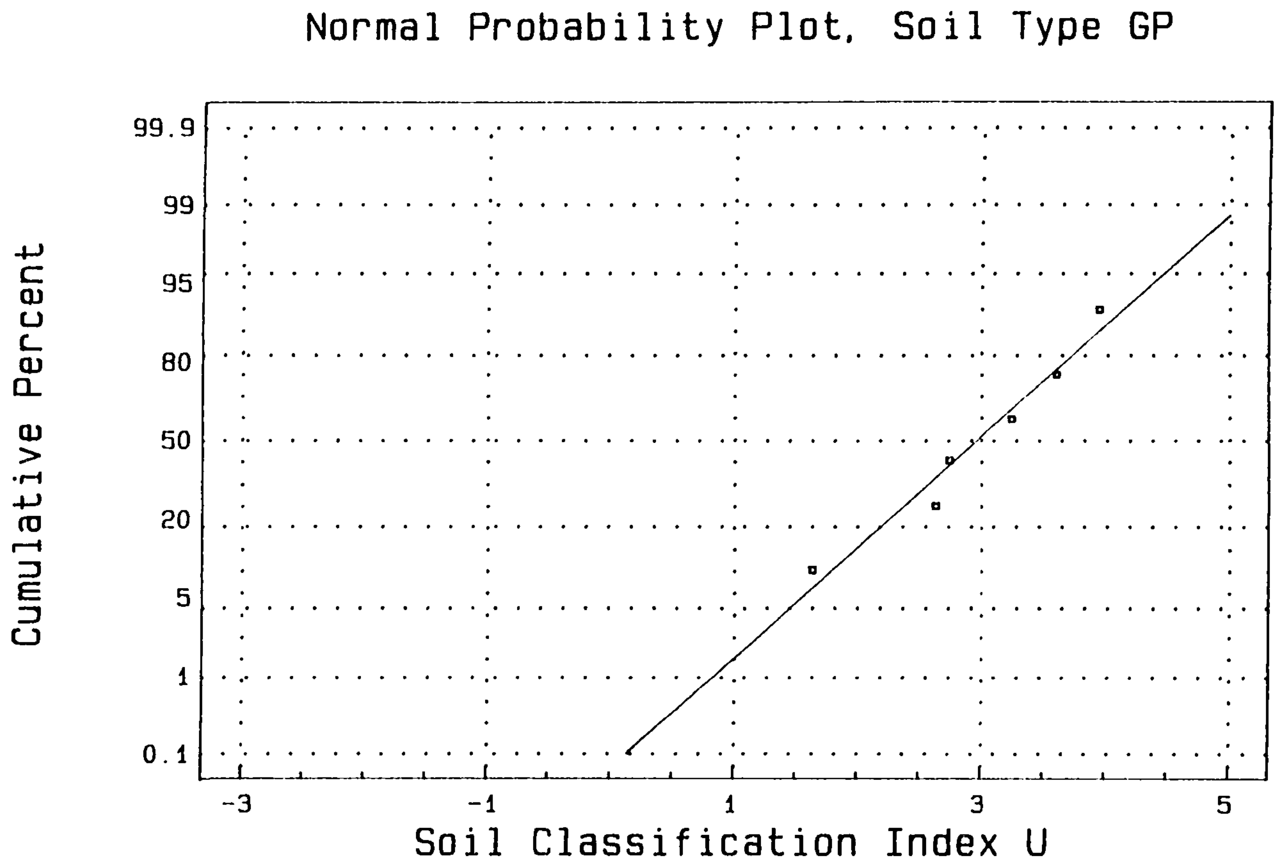


Figure 4.1 Normal Probability Plot for Soil Type GP.

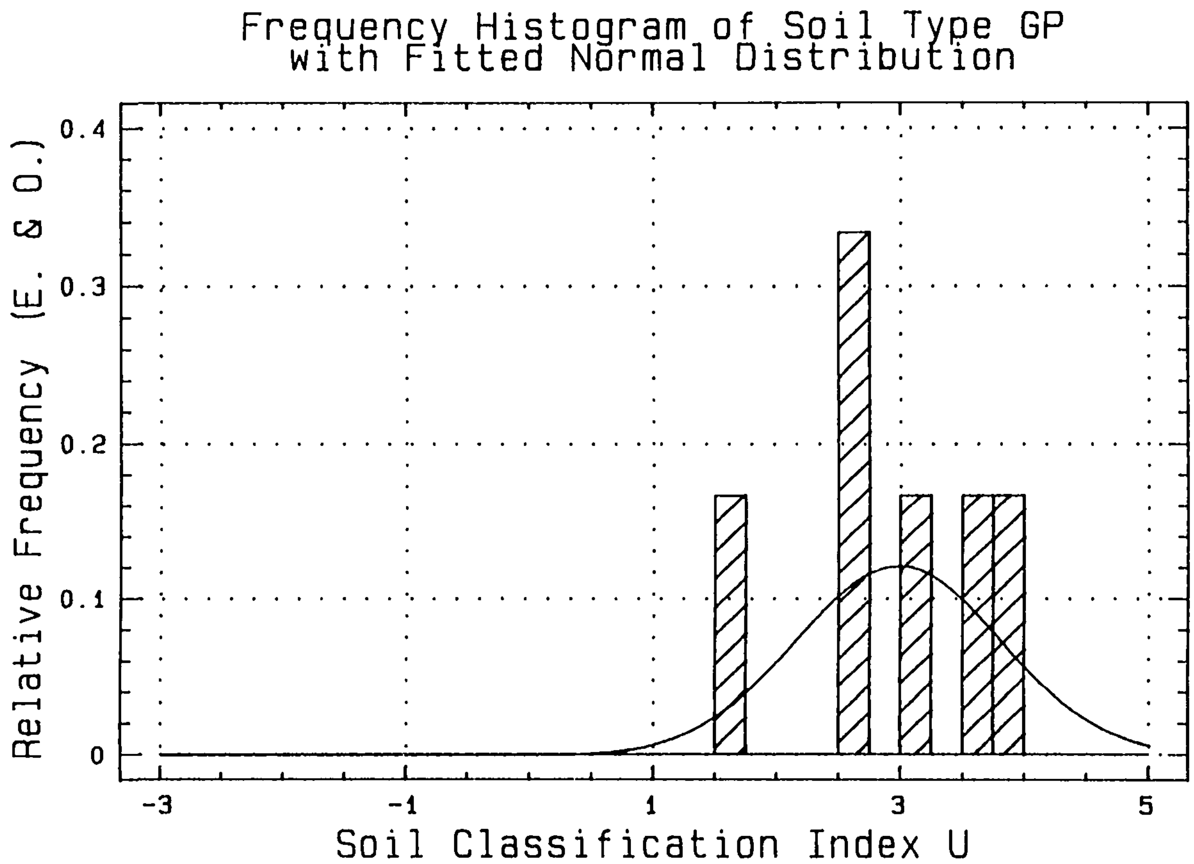


Figure 4.2 Frequency Histogram for Soil Type GP.

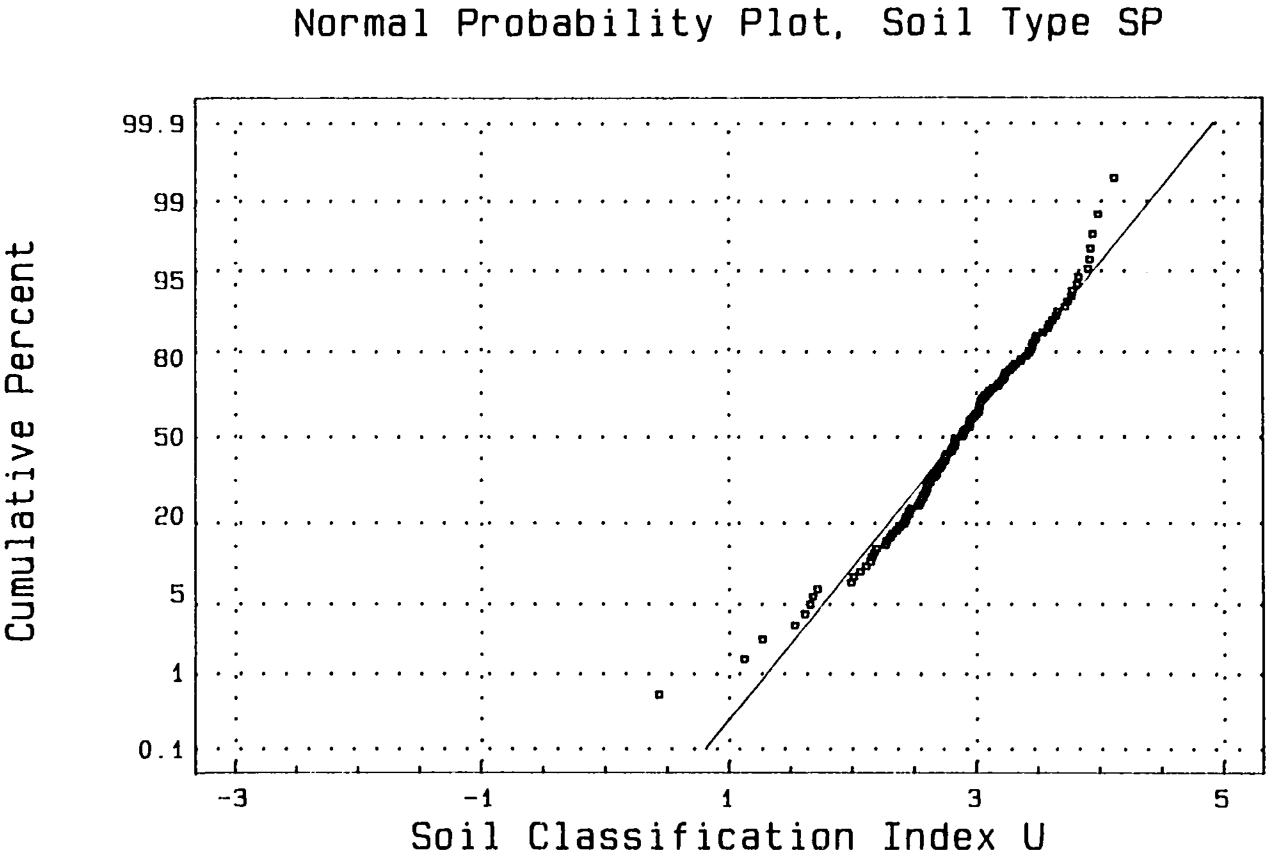


Figure 4.3 Normal Probability Plot for Soil Type SP.

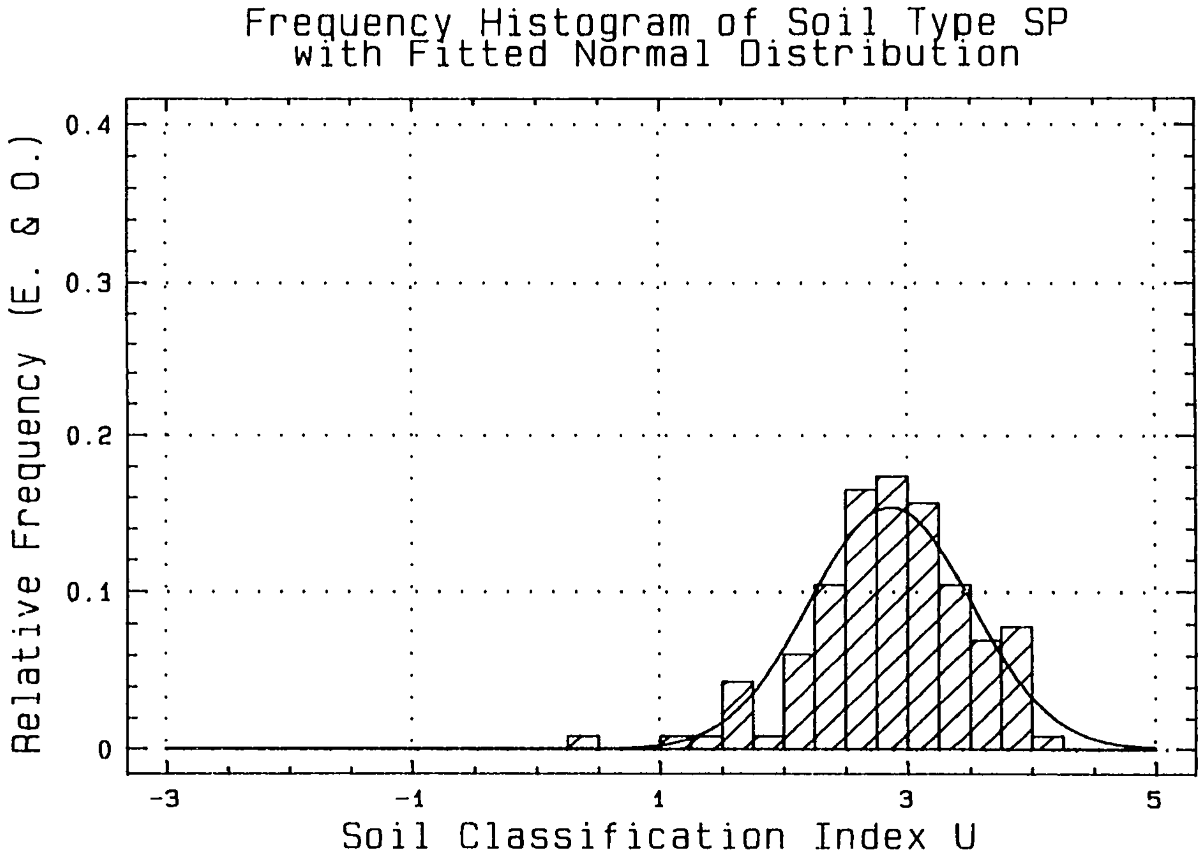


Figure 4.4 Frequency Histogram for Soil Type SP.

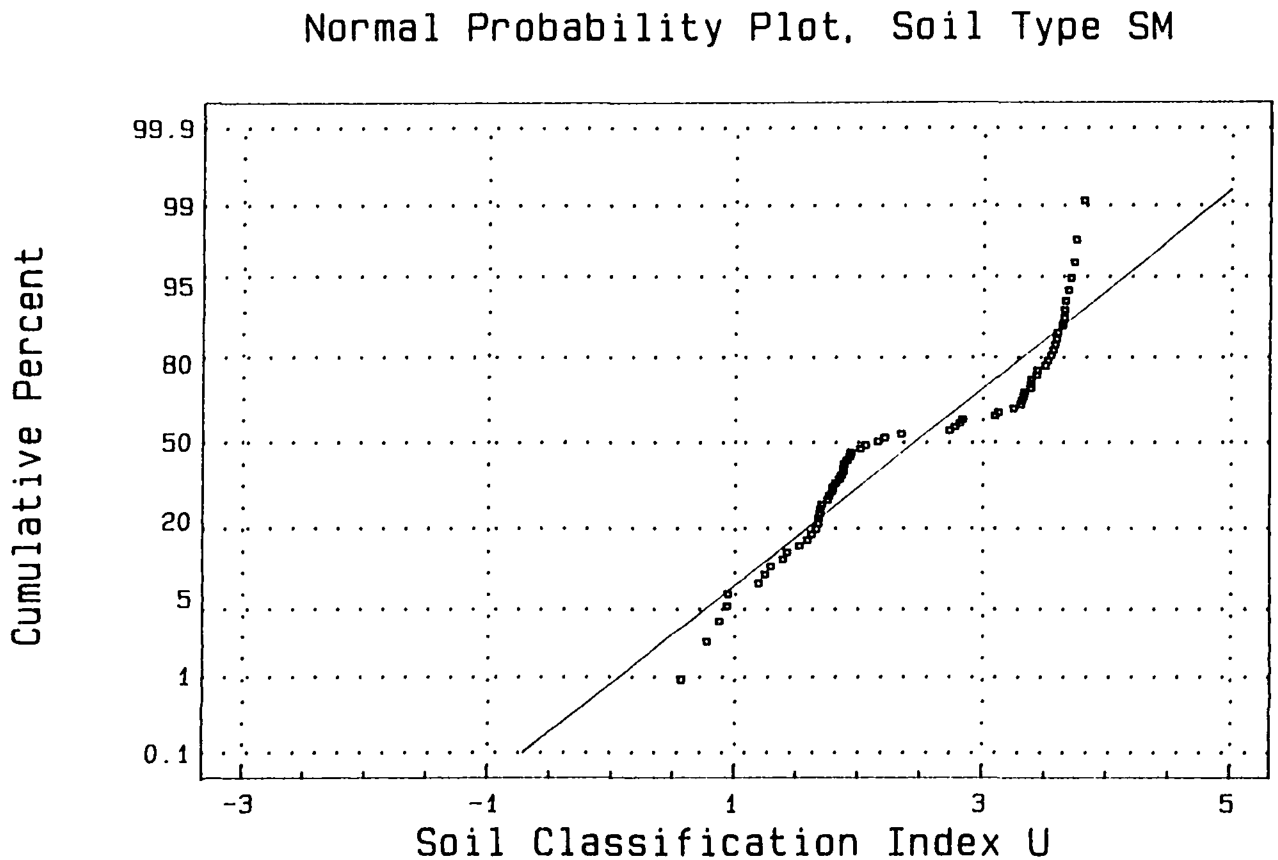


Figure 4.5 Normal Probability Plot for Soil Type SM.

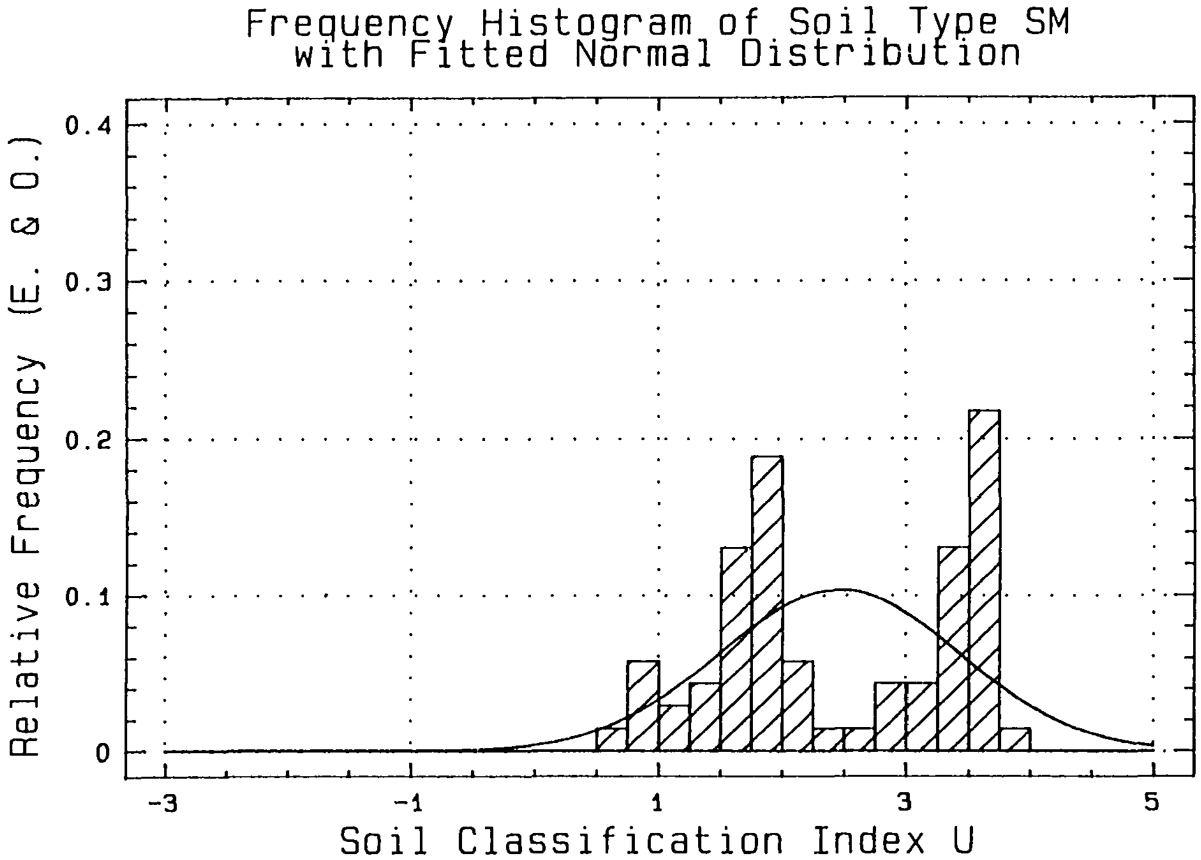


Figure 4.6 Frequency Histogram for Soil Type SM.

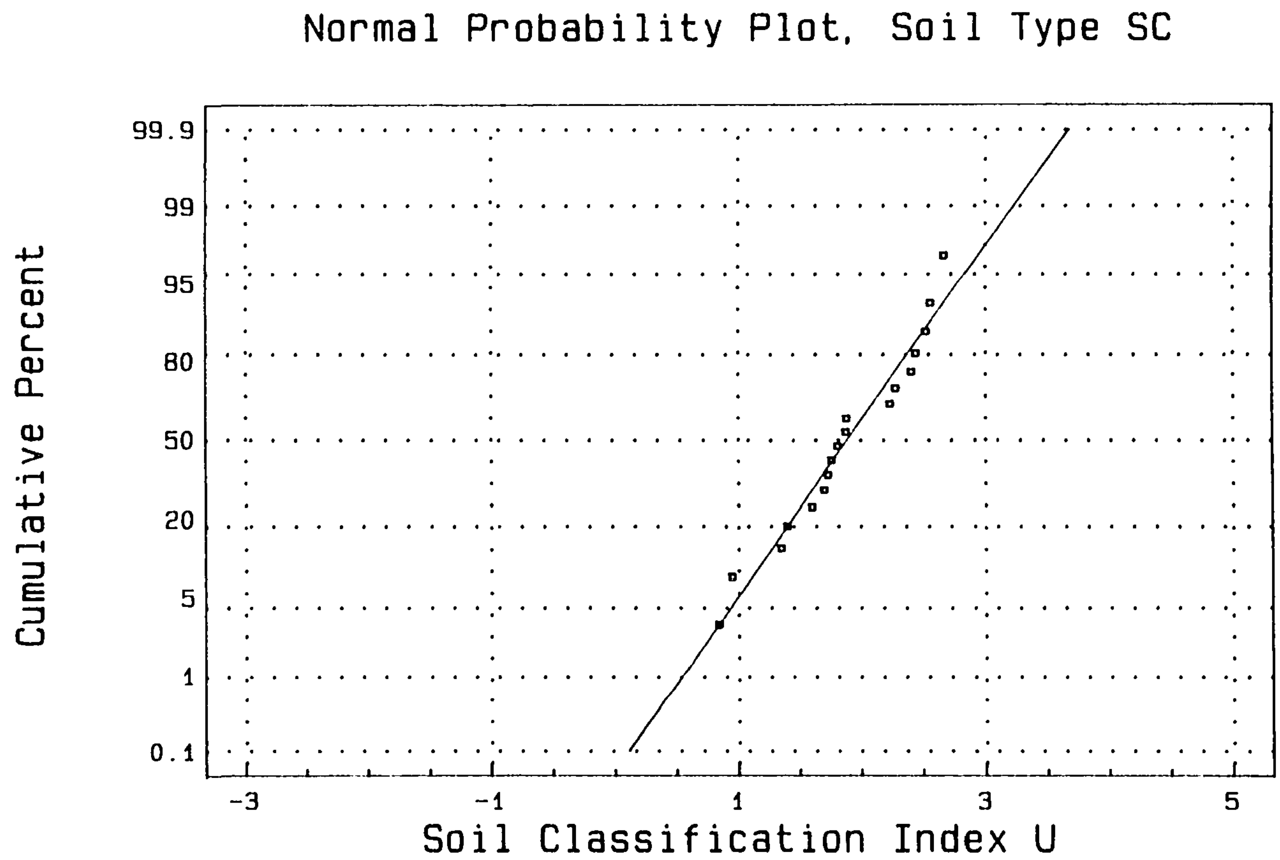


Figure 4.7 Normal Probability Plot for Soil Type SC.

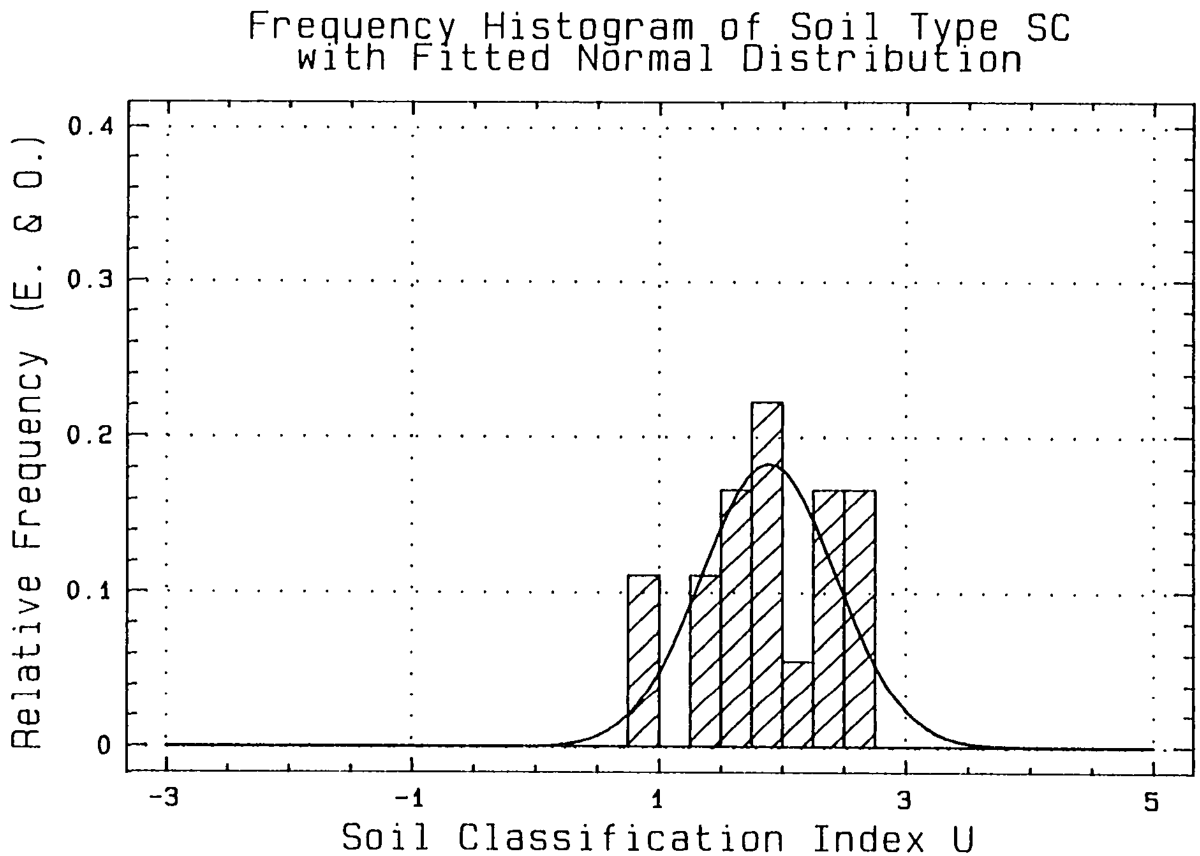


Figure 4.8 Frequency Histogram for Soil Type SC.

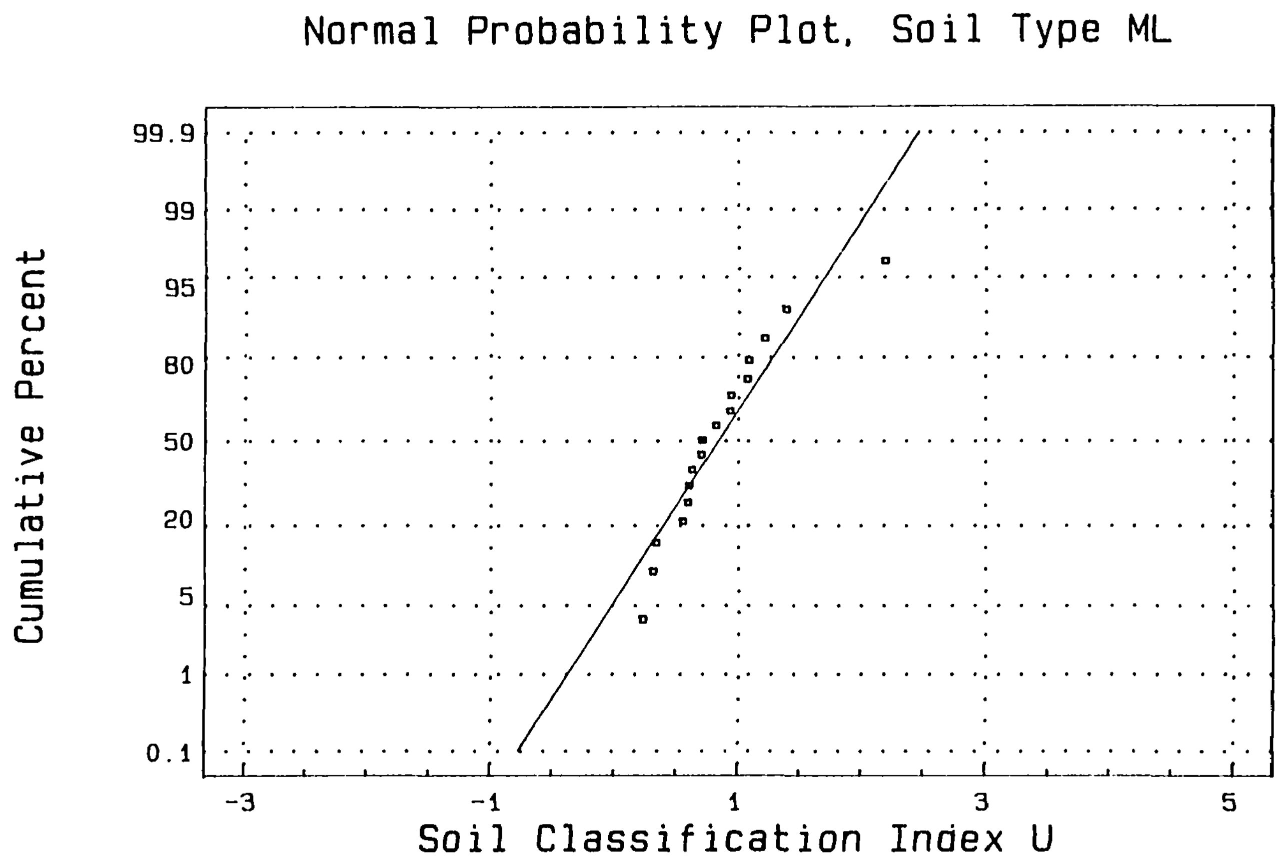


Figure 4.9 Normal Probability Plot for Soil Type ML.

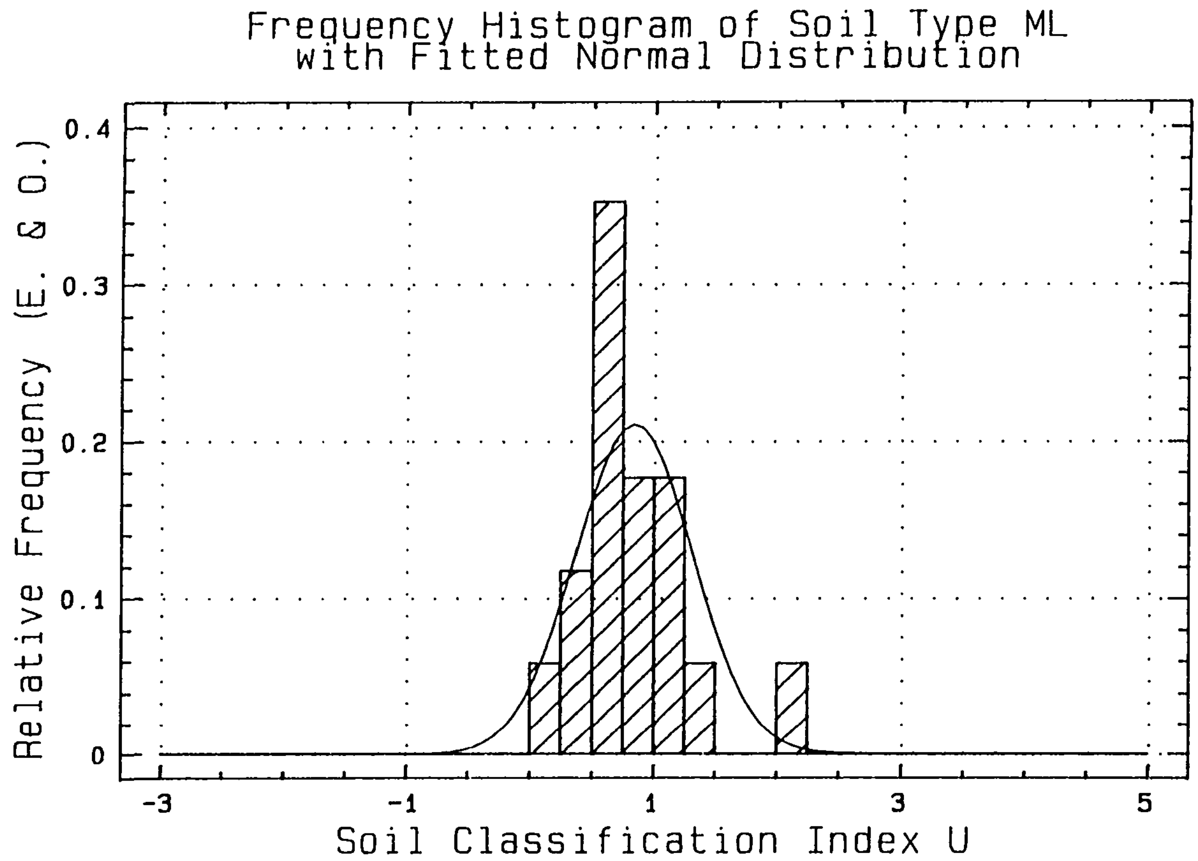


Figure 4.10 Frequency Histogram for Soil Type ML.

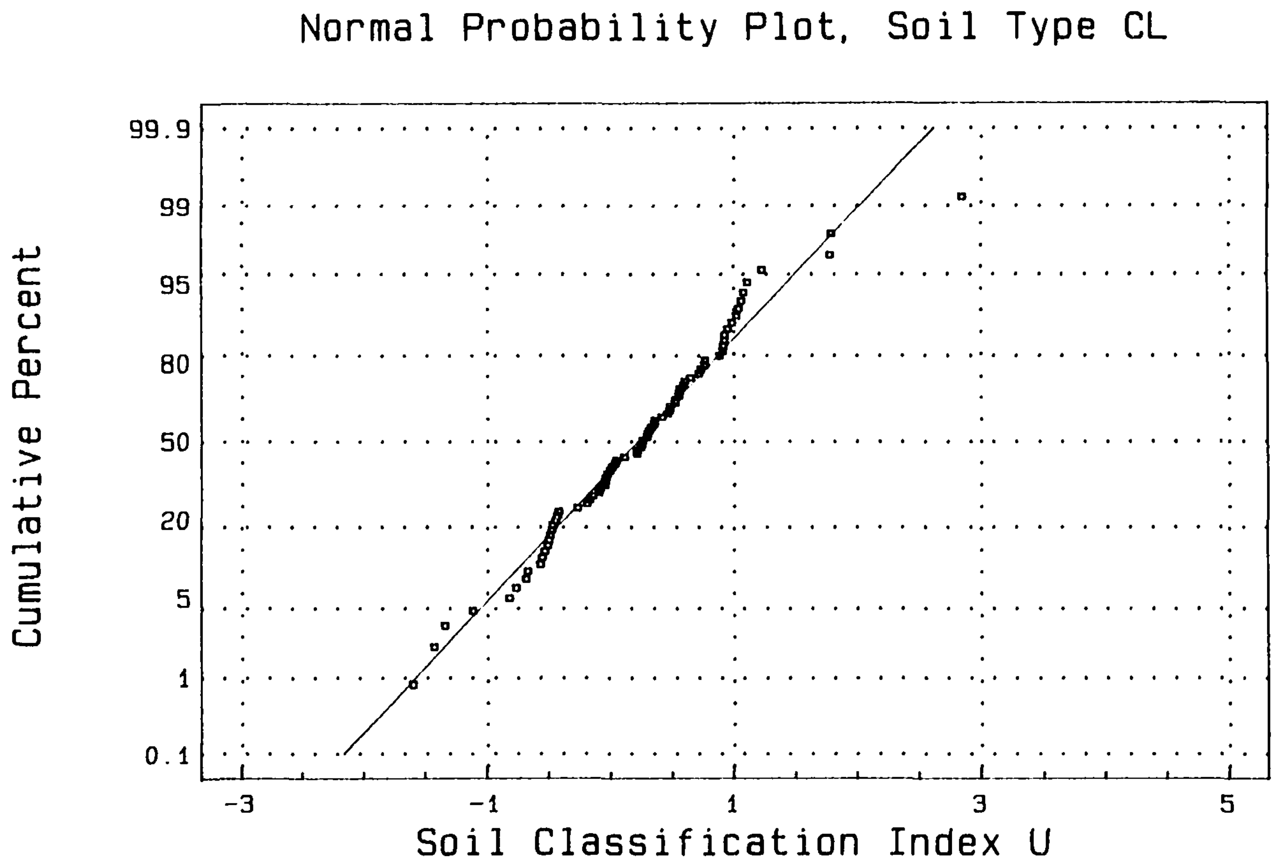


Figure 4.11 Normal Probability Plot for Soil Type CL.

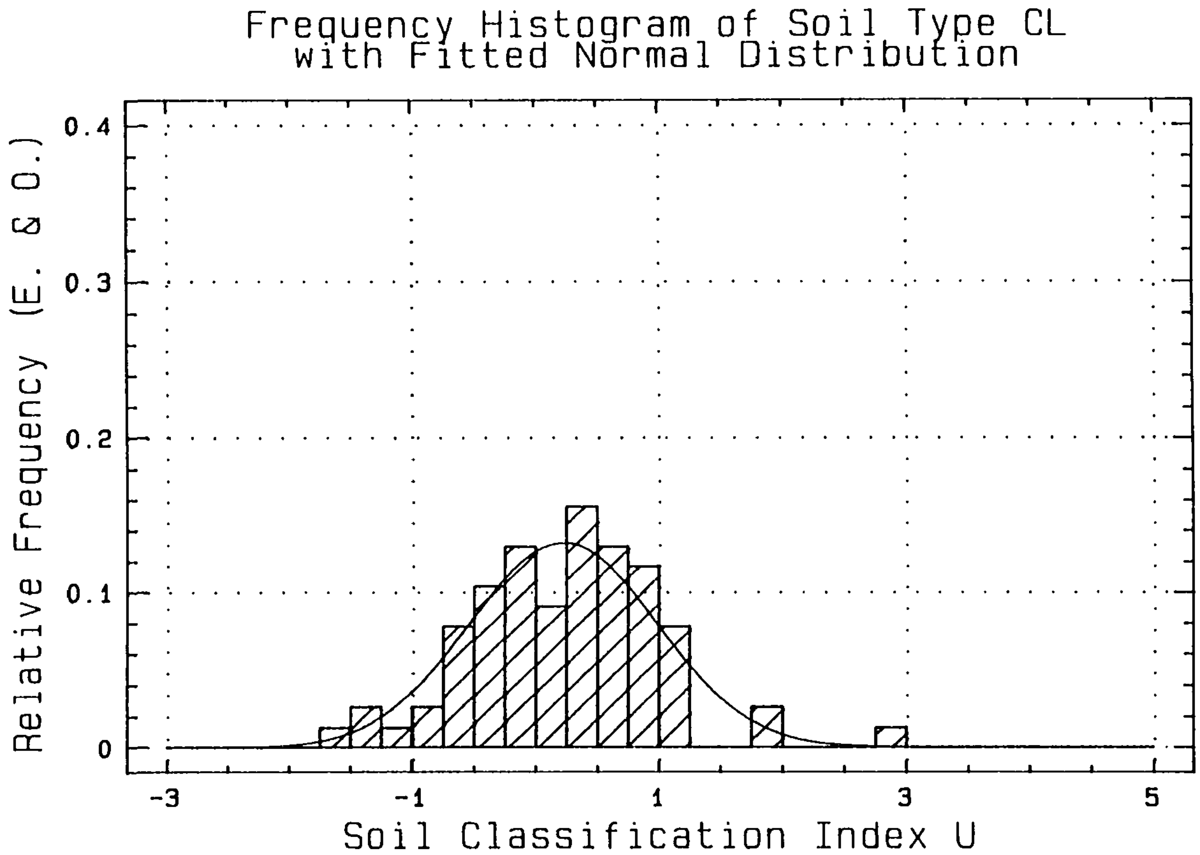


Figure 4.12 Frequency Histogram for Soil Type CL.

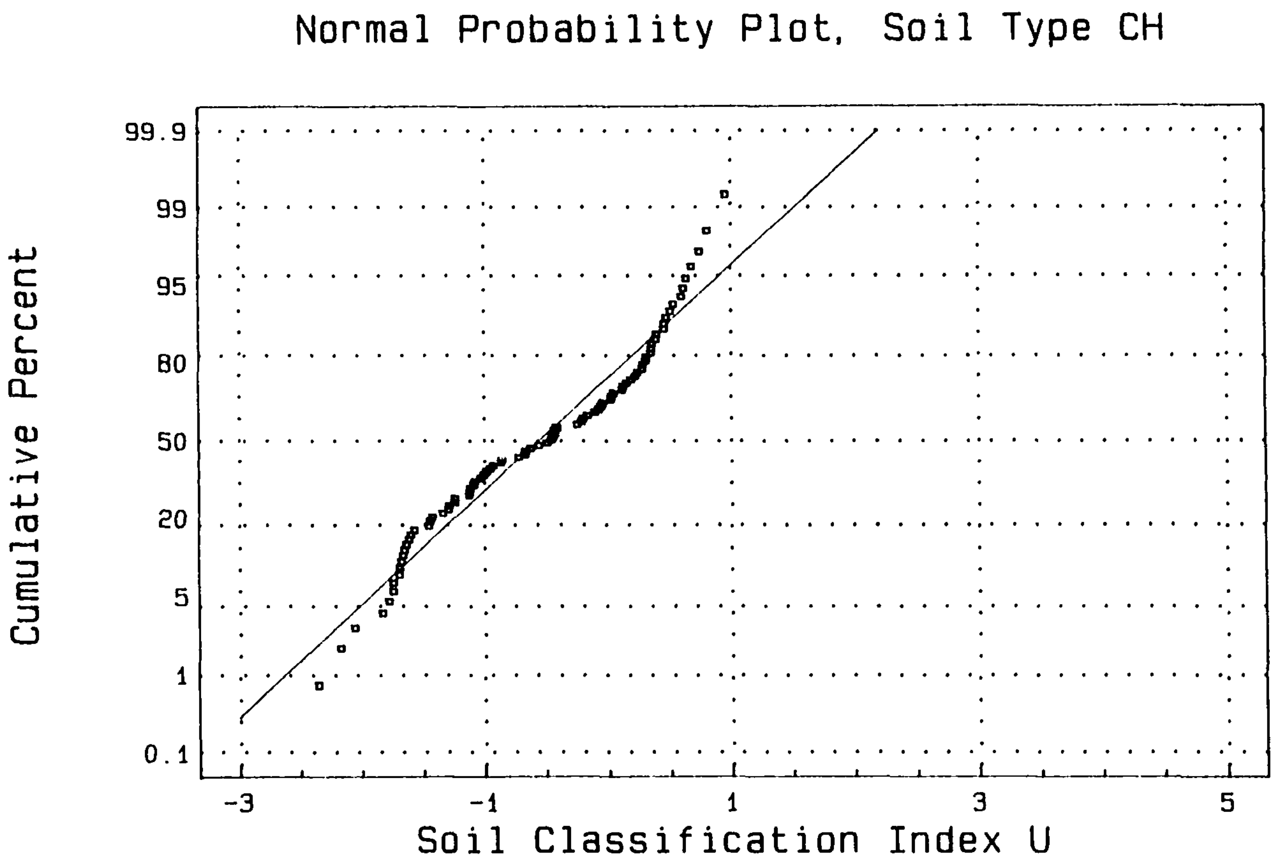


Figure 4.13 Normal Probability Plot for Soil Type CH.

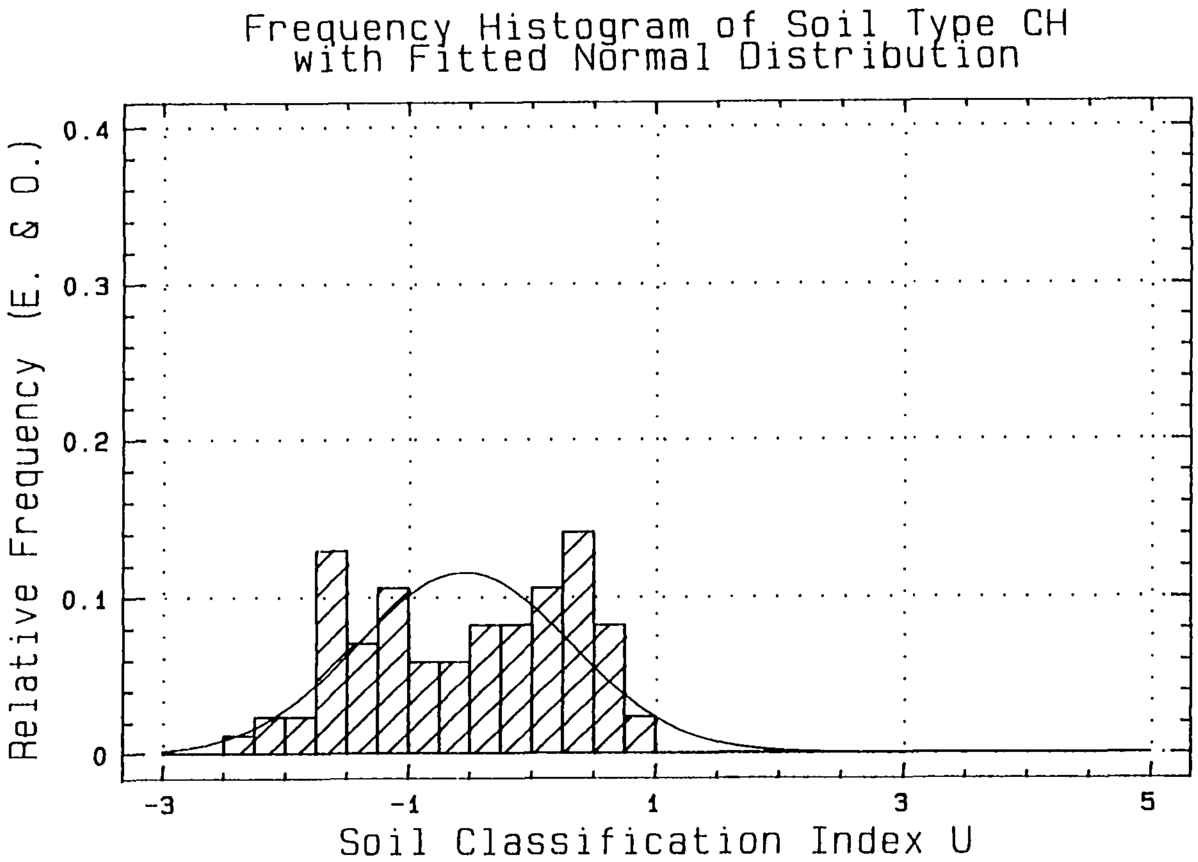


Figure 4.14 Frequency Histogram for Soil Type CH.

Table 4.1. Results of Distribution Fitting Tests

Soil Type	Sample Size	Distri.	mean	S. D.	Chi-Square Test			K - S Test		
					Est.	D. F.	S. L.	Dplus	Dmi	S. L
GP	6	F_1	2.97	0.824				0.116	0.174	0.993
SP	115	F_2	2.86	0.649	4.07	7	7.7E-1	0.035	0.065	0.708
SM	69	F_3	2.45	0.964	54.43	7	1.9E-9	0.167	0.173	0.032
SC	18	F_4	1.89	0.544				0.115	0.126	0.939
ML	17	F_5	0.85	0.472				0.141	0.097	0.889
CL	77	F_6	0.22	0.754	6.36	6	3.8E-1	0.066	0.053	0.887
CH	85	F_7	-0.55	0.861	23.90	9	4.5E-3	0.071	0.101	0.355

Note:

Distri: Distribution Function;
S. D.: Standard Deviation;
Est.: Estimate;
D.F.: Degree of Freedom;
S.L.: Significance Level;
Dplus: The maximum positive deviation of the empirical cumulative distribution over the hypothesized cumulative distribution function;
Dmi: The maximum negative deviation of the empirical cumulative distribution over the hypothesized cumulative distribution function.

K-S test results, as a more general scale of confidence, confirm the conclusions of Chi-square test. These results show that all the soil samples, except for the ones from SM and CH satisfy the normal distribution assumption with quite high confidence ($\alpha > 0.7$). Still, the confidences for CH and SM samples are low or very low ($\alpha = 0.35$ and 0.03, separately). These numeric analysis results are visualized directly in the figures from 4.1 to 4.14. It is not clear right now why CH and SM samples do not follow a normal distribution well. More **CPT** data are needed to examine and explain this phenomenon and to verify and validate the distributions of soil classification index, U, for other soil types. For the time being, the assumption of normal distribution is accepted with the significance level of 0.03 and then a prototype of **CPT** soil classification, which is better than none, could be worked out based upon these normal distributions. Figure 4.15 exhibits all the normal density functions for the seven types of soils.

4.3. Region Estimation

Region estimation is an approach similar to conventional soil classifications to classify soils. The predicted soil type for a given soil sample in this approach will depend on the regions of soil behavior units in which the corresponding in-situ behavior unit of the sample falls. The difference between the new and conventional ones is that each region of soil behavior units will correspond to not just one type of soils but several ones with different probabilities. It seems the new approach will make the soil classification complicated, but it is the real life. The basic philosophy of this new approach is to try to keep and present all original information on soil types as it is and to let users make their decisions of simplification.

The new approach is developed based upon the density functions determined previously for each type of soils. Also, the following assumptions are taken:

1). The sample density functions found previously in this chapter can be taken as the real density functions for corresponding soil types.

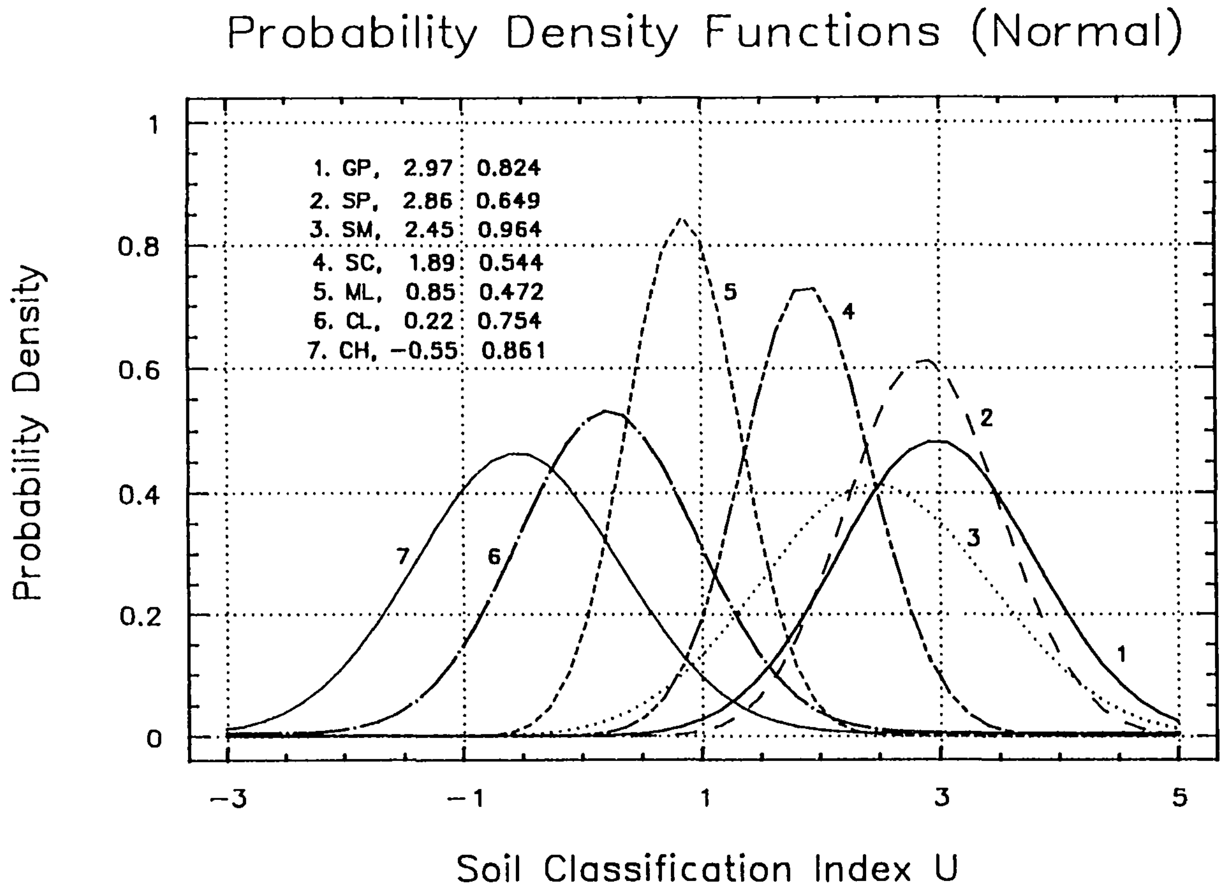


Figure 4.15 Normal Density Functions for Seven Soil Types.

2). All seven soil types shown in Table 4.1 are equally important when the determination of boundary points between adjacent soil types are considered and performed.

Consequently, boundary values $U_{i,i+1}$ for adjacent soil types, i, and, i+1, could be figured out by the conditions:

$$1 - F_i(U_{i,i+1}) = F_{i+1}(U_{i,i+1}) \quad (4.1)$$

here F_i and F_{i+1} are the cumulative distribution functions for soil type i and i+1 and i = 1, 2, ..., 6. It can be proved (see Appendix C) that the condition given by Equation (4.1) will give

$$U_{i,i+1} = \frac{\mu_{i+1}\sigma_i + \mu_i\sigma_{i+1}}{\sigma_i + \sigma_{i+1}} \quad (4.2)$$

Table 4.2. Division of Seven Regions over U Axis

Regions	Boundary Value $U_{i,i+1}$	$1 - F_i(U_{i,i+1})$
R_1 and R_2	2.91	0.471
R_2 and R_3	2.70	0.401
R_3 and R_4	2.01	0.354
R_4 and R_5	1.33	0.152
R_5 and R_6	0.61	0.306
R_6 and R_7	-0.14	0.316

The resulting boundary values are calculated according to the values given in Table 4.1 and shown in Table 4.2. These values will divide the axis of soil classification index, U, into seven regions: R_1, R_2, ... R_7. Obviously, the sizes of these regions are different, which reflects one characteristic of those distribution functions. These regions together consist of the new **CPT** soil classification. The corresponding probabilities with which each region is related to different types of soils can be determined and explained as follows.

Table 4.3 gives out the probabilities $q_{i,j}$ with which soil type i falls into region j of the U axis. These $q_{i,j}$ values are calculated by

$$q_{i,j} = F_i (U_{j,j-1}) - F_i (U_{j+1,j}) \qquad (4.3)$$

Here i, j = 1, 2, ..., 7, $U_{1,0} = +\infty$, $U_{8,7} = -\infty$, and

$$\sum_{j=1}^{7} q_{i,j} = 1 \qquad (4.4)$$

Table 4.3. Probabilities with Which Each Type of Soil Falls in Each Region

Regions	R_1	R_2	R_3	R_4	R_5	R_6	R_7
Distribu-tion	Probability $q_{i,j}$ over Regions						
F_1	0.5285	0.0991	0.2274	0.1214	0.0214	0.0021	0.0000
F_2	0.4715	0.1274	0.2821	0.1098	0.0090	0.0002	0.0000
F_3	0.3200	0.0811	0.2452	0.2323	0.0940	0.0238	0.0036
F_4	0.0308	0.0380	0.2848	0.4944	0.1428	0.0090	0.0001
F_5	0.0000	0.0000	0.0040	0.1478	0.5421	0.2873	0.0187
F_6	0.0002	0.0003	0.0060	0.0642	0.2353	0.3781	0.3159
F_7	0.0000	0.0000	0.0009	0.0134	0.0754	0.2261	0.6841
SubTotal	1.3510	0.3459	1.0504	1.1833	1.1200	0.9266	1.0224

Table 4.4. Probability with Which Each Region Receives Each Type of Soil

Region	R_1	R_2	R_3	R_4	R_5	R_6	R_7
Soil Type	Probability $p_{i,j}$ over Regions						
GP	0.3912	0.2865	0.2165	0.1026	0.0191	0.0023	0.0000
SP	0.3490	0.3683	0.2686	0.0928	0.0080	0.0002	0.0000
SM	0.2369	0.2345	0.2334	0.1963	0.0839	0.0257	0.0035
SC	0.0229	0.1098	0.2711	0.4178	0.1275	0.0097	0.0001
ML	0.0000	0.0000	0.0038	0.1249	0.4840	0.3100	0.0183
CL	0.0000	0.0009	0.0057	0.0542	0.2100	0.4080	0.3090
CH	0.0000	0.0000	0.0009	0.0113	0.0675	0.2441	0.6691

Table 4.4 presents the probabilities $p_{i,j}$ with which region j will receive soil type i and

$$p_{i,j} = \frac{q_{i,j}}{\sum_{i=1}^{7} q_{i,j}} \tag{4.5}$$

here, i, j = 1, 2, ..., 7, too. These values quantitatively describe the overlaps among different types of soils. They indicate that a region can correspond to several types of soils with different probabilities. For example, region R_1 corresponds to soil GP with probability of 0.39, soil SP with 0.35, soil SM with 0.24, and soil SC with 0.02. Also, the degree of the overlaps in Table 4.4 can be reduced by grouping some types of soils, as shown in Table 4.5.

Table 4.5. Simplified Results from Table 4.4

Region	R_1	R_2	R_3	R_4	R_5	R_6	R_7
Soil Type	Probability $p_{i,j}$ over Regions						
GP, SP, SM	0.9771	0.8893	0.7185	0.3971	0.1110	0.0282	0.0035
SC, ML	0.0229	0.1098	0.2749	0.5427	0.6115	0.3197	0.0184
CL, CH	0.0000	0.0009	0.0066	0.0655	0.2775	0.6521	0.9781
GP, SP, SM, SC	1.0000	0.9991	0.9896	0.8096	0.2358	0.0379	0.0036
ML, CL, CH	0.0000	0.0009	0.0104	0.1904	0.7615	0.9621	0.9964

Table 4.5 indicates that the sandy soils (GP, SP, SM) generally fall in Region 1, 2, and 3, the silty soils (SC, MC) in Region 4 and 5, and the clayey soils (CL, CH) in Region 6 and 7. If the silty soils are further divided and merged with the sandy and clayey soils, as shown at the bottom of Table 4.5, the boundary value between Region 4 and 5 ($U_{4,5}$ = 1.33, given in Table 4.2) can reasonably be taken as the dividing point. These results are consistent quite well with the ones from the empirical analysis performed previously in Chapter 3. The advantage here is that quantitative description on overlaps are also provided. Now, all the results for Region Estimation are summarized and visualized in Figure 4.16. The vertical relative lengths in this chart represent the values of the probabilities $p_{i,j}$. If the **CPT** behavior unit of a soil sample is given, the corresponding soil type can be predicted accordingly.

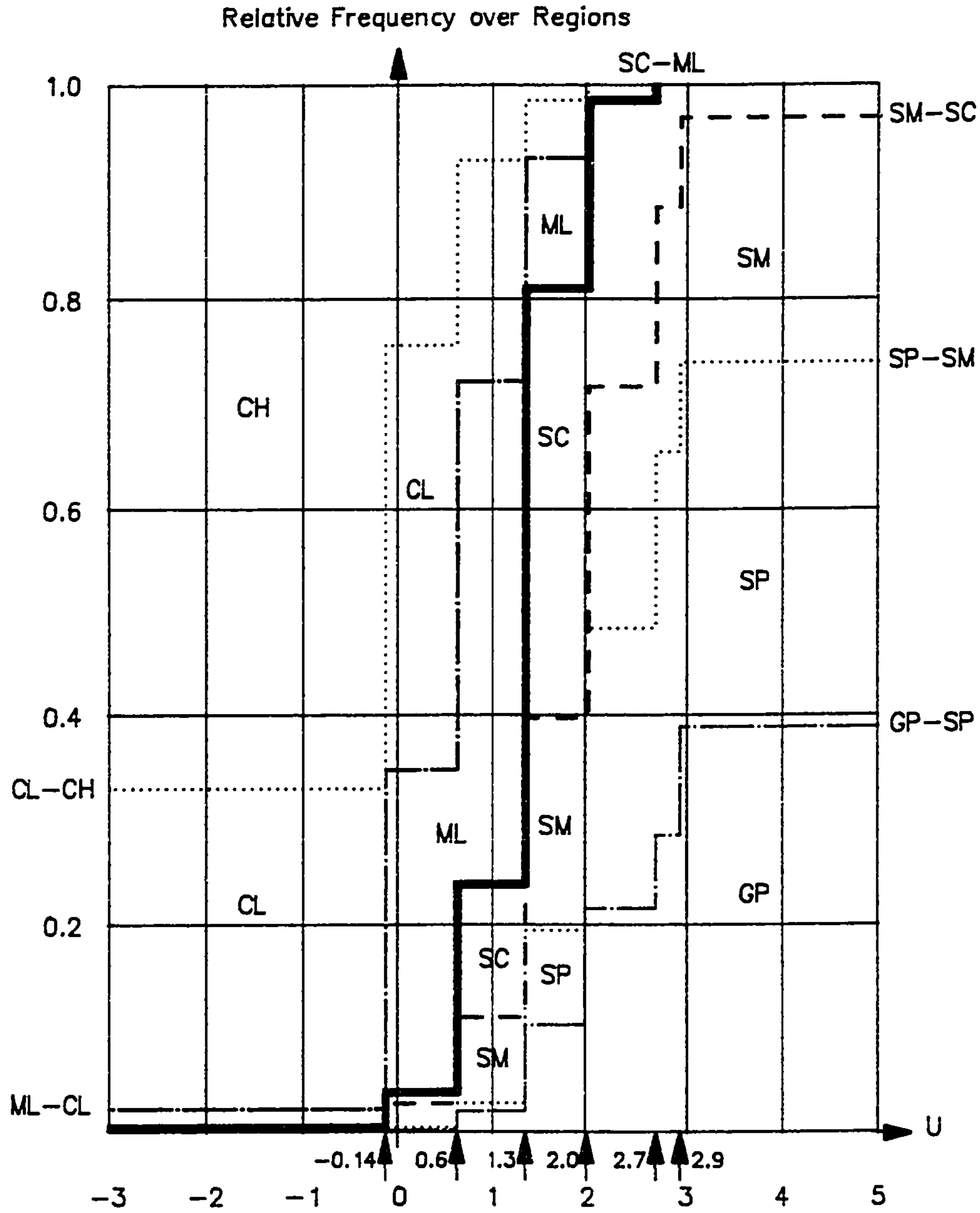


Figure 4.16 Region Estimation Chart for CPT Soil Classification.

Region Estimation has a problem of over simplification on real situation. This is because in this approach different points of a region are supposed to have the exactly same statistical property so that they are treated in an exactly same way. Unfortunately, this assumption is not true in most cases since different points in a region will have different probabilities to correspond to (receive) different types of soils. In some cases, this kind of difference is quite large. Therefore, from a theoretical point of view, the accuracy provided by this approach is reduced to some extent. This problem will be solved by another method called "Point Estimation". In that new approach, every point along the U axis will be treated distinctively. Its details will be presented and discussed in the next section.

4.4. Point Estimation

Point Estimation is an approach to classify soils directly by probability. Each possible values of soil in-situ behavior unit, U, will be evaluated individually by this approach. The basic question it intends to answer is: given a specific value of U, what are the probabilities with which the corresponding soil sample belongs to different types of soils? The whole approach is based upon a probabilistic model of two dimensions, one of which is the continuous random variable of soil classification index, U, (representing soil in-situ behavior units). The other is a discrete random variable representing soil types. Both variables are defined in soil layers. The physical meaning of this probabilistic model can be understood by discussing a kind of imagined theoretical test first.

Assume that there is a box containing balls with different colors. The number of colors is N and the number of balls is infinite for each color. Then, if a ball is taken out of the box, the probability p(color) for a specific color being selected is p(color) = 1/N.

Now consider another test. In addition to the above assumptions, suppose that a kind of quantity, x, can be measured from each ball and that this quantity is a continuous random variable which follows a different specific distribution for each color. Thus, if a ball is again taken out of that box and the corresponding x value is measured from that ball, the probability for a specific color being selected will be modified by the information of the x value. This is because different color balls will give this x value with different probabilities. Then, what will be the modification by this x value on the probability with which a ball with a specific color is selected or $p(\text{color} \mid X{=}x) = ?$

Obviously, this is a problem of two dimensional random variables, one dimension being the random variable, x, with continuous distribution and the other being the discrete random variable, color. The solution to this problem, according to the Multiplication Rule in Probability Theory, will be

$$P(\, color \mid X{=}x\,) = \frac{P(\,(\,X{=}x\,) \cap (\,color\,)\,)}{P(\,X{=}x\,)} \qquad (4.6)$$

Since X is a continuous random variable, $P(X = x) = 0$ exists. So the above equation has to be replaced by

$$P(color|X=x) = \lim_{\Delta x \to 0} \frac{P((x \leq X < x+\Delta x) \cap (color))}{P(x \leq X < x+\Delta x)} \quad (4.7)$$

This replacement is valid since the measurement of the random variable X can not be exact due to a possible measurement error.

Now, imagine that the big box is the earth and the countless balls are different soil layers. The continuous random variable X is the corresponding soil in-situ behavior unit represented by index U and the different colors stand for different types of soils we want to identify in a soil classification. Then, the previous analysis on the problem of statistically anticipating a ball color according to an X value can be applied here to statistically predict a soil type based on a single observation of soil in-situ behavior unit, U. That is

$$P(soiltype|U=u) = \lim_{\Delta u \to 0} \frac{P((u \leq U < u+\Delta u) \cap (soiltype))}{P(u \leq U < u+\Delta u)} \quad (4.8)$$

If the soil type is represented by a discrete numeric random variable SI, Equation (4.8) will be

$$P(SI=si|U=u) = \lim_{\Delta u \to 0} \frac{P((u \leq U < u+\Delta u) \cap (SI=si))}{P(u \leq U < u+\Delta u)} \quad (4.9)$$

here SI means Soil type Index and will take values of 1, 2, ... M. M is the number of soil types in a concerned soil classification. Also, the uppercase of SI or U means a variable and the lowercase is the value of that variable.

Suppose that F(u, si) is the two dimensional probability distribution function of index U and SI and $F_m(u)$ is its marginal distribution function of U. It is known by definition that the probabilities of events in Equation (4.9) can be rewritten in terms of F(u, si) and its marginal distribution function $F_m(u)$. That is

$$P(\,(\,u \le U < u+\Delta u\,) \cap (\,SI = si\,)\,) \;=\; F(\,u+\Delta u, si\,) \;-$$

$$- \,F(\,u, si\,) - F(\,u+\Delta u, si-1\,) + F(\,u, si-1\,) \qquad (4.10)$$

and

$$P(\,u \le U < u+\Delta u\,) \;=\; F_m(\,u+\Delta u\,) - F_m(\,u\,) \qquad (4.11)$$

Now, the foregoing problem has become how to decide the two dimensional probability distribution function, F(u, si).

In Probability Theory, the two dimensional distribution function F(u,si) is defined as

$$F(\,u, si\,) \;=\; \sum_{y=1}^{si} \int_{-\infty}^{u} f(\,x, y\,)\,dx \qquad (4.12)$$

or

$$F(\,u, si\,) \;=\; \int_{-\infty}^{u} \sum_{y=1}^{si} f(\,x, y\,)\,dx \qquad (4.13)$$

here, x means u and y stands for si. x and y are used in place of u and si in order to observe the rule of integration and avoid other possible confusions. This practice will be followed in the entire process of derivation in this section.

The corresponding marginal distribution function of U is then

$$F_m(u) = \int_{-\infty}^{u} \sum_{si=1}^{M} f(x, si) \, dx \tag{4.14}$$

or

$$F_m(u) = \sum_{si=1}^{M} \int_{-\infty}^{u} f(x, si) \, dx \tag{4.15}$$

here, M is the number of soil types in the concerned soil classification. Also, the marginal distribution function of SI has the form

$$F_m(si) = \sum_{y=1}^{si} \int_{-\infty}^{+\infty} f(u, y) \, du = \sum_{y=1}^{si} q(y) \tag{4.16}$$

where

$$q(si) = q(y) = \int_{-\infty}^{+\infty} f(u, si) \, du \tag{4.17}$$

Due to the basic properties of summation and integration, Equation (4.12) is equivalent to Equation (4.13) and Equation (4.14) is equivalent to Equation (4.15).

Now, we have to determine the two dimensional density function f(u, si) in above equations. According to the definition of F(u, si), we have

$$F(u, si) = P(U \le u, SI \le si) \tag{4.18}$$

$$F(u, si-1) = P(U \le u, SI \le si-1) \tag{4.19}$$

Subtract Equation (4.19) from Equation (4.18), we get

$$F(u,si) - F(u,si\text{-}1) = P(U\le u, SI\le si) - P(U\le u, SI\le si\text{-}1)$$

$$= P(\,(\,U\le u\,) \cap (\,SI=si\,)\,) \qquad (4.20)$$

since SI is a discrete random variable. Therefore,

$$P(\,(\,U\le u\,) \cap (SI=si\,)\,) = F(\,u, si\,) - F(\,u, si\text{-}1\,)$$

$$= \sum_{y=1}^{si} \int_{-\infty}^{u} f(\,x,y\,)\;dx - \sum_{y=1}^{si-1} \int_{-\infty}^{u} f(\,x,y\,)\;dx$$

$$= \int_{-\infty}^{u} f(\,x,\;si\,)\;dx \qquad (4.21)$$

On the other hand, the Multiplication Rule will give us

$$P(\,(U\le u) \cap (SI=si)\,) = P(SI=si)\,P(\,(U\le u)\,|\,(SI=si)\,) \qquad (4.22)$$

Here

$$P(\,(\,U\le u\,)\,|\,(\,SI=si\,)\,) = \int_{-\infty}^{u} g_{si}(\,x\,)\;dx \qquad (4.23)$$

The $g_{si}(u) = g_{si}(x)$ in Equation (4.23) is the conditional density function of index U for soil type si. Also, according to Equation (4.16), P(SI=si) in Equation (4.22) is

$$P(\ SI=si\ ) = F_m(\ si\ ) - F_m(\ si-1\ ) = q(si) \tag{4.24}$$

This is the probability with which a certain type of soil will be met in general if no soil in-situ behavior information is available. This probability, written as q(si), should be a constant for soil type si.

Now, substitute Equation (4.23) and (4.24) into Equation (4.22) and then Equation (4.22) into Equation (4.21), we get

$$\int_{-\infty}^{u} f(x,\,si)\ dx = q(si) \int_{-\infty}^{u} g_{si}(x)\ dx \tag{4.25}$$

Since u can be any real value, Equation (4.25) exists only if

$$f(u,\,si) = q(si)\ g_{si}(u) \tag{4.26}$$

This is the two dimensional density function searched.

Then, substitute Equation (4.12) into the right hand of Equation (4.10) and rearrange it. We have

$$P(\ (\ u \le U < u+\Delta u\ ) \cap (\ SI=si\ )\ ) =$$

$$= \sum_{y=1}^{si} \left(\int_{-\infty}^{u+\Delta u} f(x,y)\ dx - \int_{-\infty}^{u} f(x,y)\ dx \right) -$$

$$- \sum_{y=1}^{si-1} \left(\int_{-\infty}^{u+\Delta u} f(x,y)\ dx - \int_{-\infty}^{u} f(x,y)\ dx \right) \tag{4.27}$$

After some derivation and using Equation (4.26), we get

$$P(\ (\ u \le U < u+\Delta u\ ) \cap (\ SI = si\ )\ ) =$$

$$= \int_u^{u+\Delta u} f(x,\ si)\ dx = q(si) \int_u^{u+\Delta u} g_{si}(x)\ dx \qquad (4.28)$$

Also, plug Equation (4.15) into Equation (4.11) and use Equation (4.26) again. We can obtain

$$P(\ u \le U < u+\Delta u\ ) = \sum_{si=1}^{M} \int_u^{u+\Delta u} f(\ x, si\ )\ dx$$

$$= \sum_{si=1}^{M} q(si) \int_u^{u+\Delta u} g_{si}(\ x\ )\ dx \qquad (4.29)$$

Therefore, the conditional probability determined by Equation (4.9) will be

$$P(SI=si \mid U=u) = \lim_{\Delta u \to 0} \frac{q(si) \int_u^{u+\Delta u} g_{si}(x)\ dx}{\sum_{si=1}^{M} q(si) \int_u^{u+\Delta u} g_{si}(x)\ dx} \qquad (4.30)$$

Since Δu is an infinitesimal quantity, Equation (4.30) can be rewritten as

$$P(SI=si \mid U=u) = \lim_{\Delta u \to 0} \frac{q(si)\ g_{si}(\ u\ )\ \Delta u}{\sum_{si=1}^{M} q(si)\, g_{si}(\ u\ )\, \Delta u} \qquad (4.31)$$

The limit in Equation (4.31) is actually independent of Δu. Consequently, we have

$$P(\ SI=si\,|\,U=u\ ) = \frac{q(si)\ g_{si}(\ u\ )}{\sum_{si=1}^{M} q(si)\ g_{si}(\ u\ )} \tag{4.32}$$

This is the basic formula to perform the point estimation.

It can be seen from Equation (4.32) that a series of conditional density functions $g_{si}(u)$, si = 1, 2, ..., M, have to be determined first before that formula can be practically used. Fortunately, we have done this job in Section 4.2, where we derived the conditional probability distributions. The condition in that situation is the given soil types. As defined previously, these soil types are represented by the variable SI and its values correspond to soil types in following order: GP, SP, SM, SC, ML, CL, CH. So what we need to do next is to decide q(si), the marginal mass function of SI.

There are two possible ways to determine or estimate q(si). The first one is based upon a subjective judgement. Back to the theoretical test discussed at beginning of this section. Under the described conditions of test, if a ball is taken randomly out of the box, we know that the probability for a certain color being selected is equal to 1/N. N is the number of colors in that box. Our soil classification problem can be considered in a similar way. Then, what is our basic statistic belief of the frequencies with which various soil types are encountered in site investigations? Is there any bias in the current **CPT** database? Unfortunately, due to the limited data available, these questions can not be answered with full confidence at this time. As a subjective prediction or an approximation, it is secure to take q(si) = 1/M, si = 1, 2, ..., M. M is the number of soil types in a related soil classification. In our case, M = 7.

The second way to determine q(si) relies on a statistical technique called the "Goodness of Fit Test". It is more directly based upon cone testing data available. In the method, the observed frequency and the expected frequency are calculated for each soil type according to testing data available and the assumed values of q(si). Then, a probability statement, indicating how well the assumed q(si) fits the testing data, can be worked out by a likelihood ratio test. The results of q(si) from this method might be different with 1/M. However, if the assumptions taken in the first method are correct and the data used in the second method is really random and reliable, it is likely that the results from these two methods will be close to each other. Due to the lack of sufficient CPT testing data, such a likelihood ratio test is not performed here. Instead, only some assumed values of q(si) are taken in order to illustrate how Point Estimation can help engineers make their estimations and predictions of soil types.

The following q(si) values have been adopted in this research for each type of soil: 1/7, 1/3, and 1/2. q(si) taking 1/7 means that the seven (7) types of soils generally have the same probability to be encountered. Some people may think that certain types of soils can have higher probabilities to be met than others in reality. The results from q(si) = 1/3 and 1/2 will provide a good scale range for their estimation purpose.

When one soil type takes A as its q(si) value, other six (6) soil types are assumed to have the value of (1-A)/6 for their corresponding q(si). Here, A stands for one of 1/7, 1/3, and 1/2. Such an assumption is taken in an attempt to make the problem simpler. It is expected that this simplification will give out average conditional probabilities for the cases where A is not equal to 1/7. The conditional probabilities obtained for seven (7) types of soils with different q(si) values have been presented from Figure 4.17 to Figure 4.23. Again, Figure 4.19 exhibits an abnormal behavior for soil type SM, which means more reliable data are needed to obtain correct results for this figure. Now, the point estimation on soil type can be performed according to these conditional probability charts if a **CPT** soil behavior unit U is given.

4.5. Summary

Two statistical approaches for predicting soil types, named as Region Estimation and Point Estimation, are suggested in this chapter to quantitatively describe the overlaps among different types of soils in a **CPT** soil classification. They are based upon the probability distributions of soil behavior unit, U, of seven (7) types of soils. These distributions are assumed to have a normal form and such an assumption has been checked by the in-situ soil behavior units of different soil types gained in Chapter 3. The corresponding coefficients of these normal distributions are also calculated according to the same data. The examinatorial results show that the assumption of normal distribution is acceptable with a high confidence ($\alpha > 0.7$) for most soil types except for SM and CH. It is not clear why soil type SM and CH do not follow a

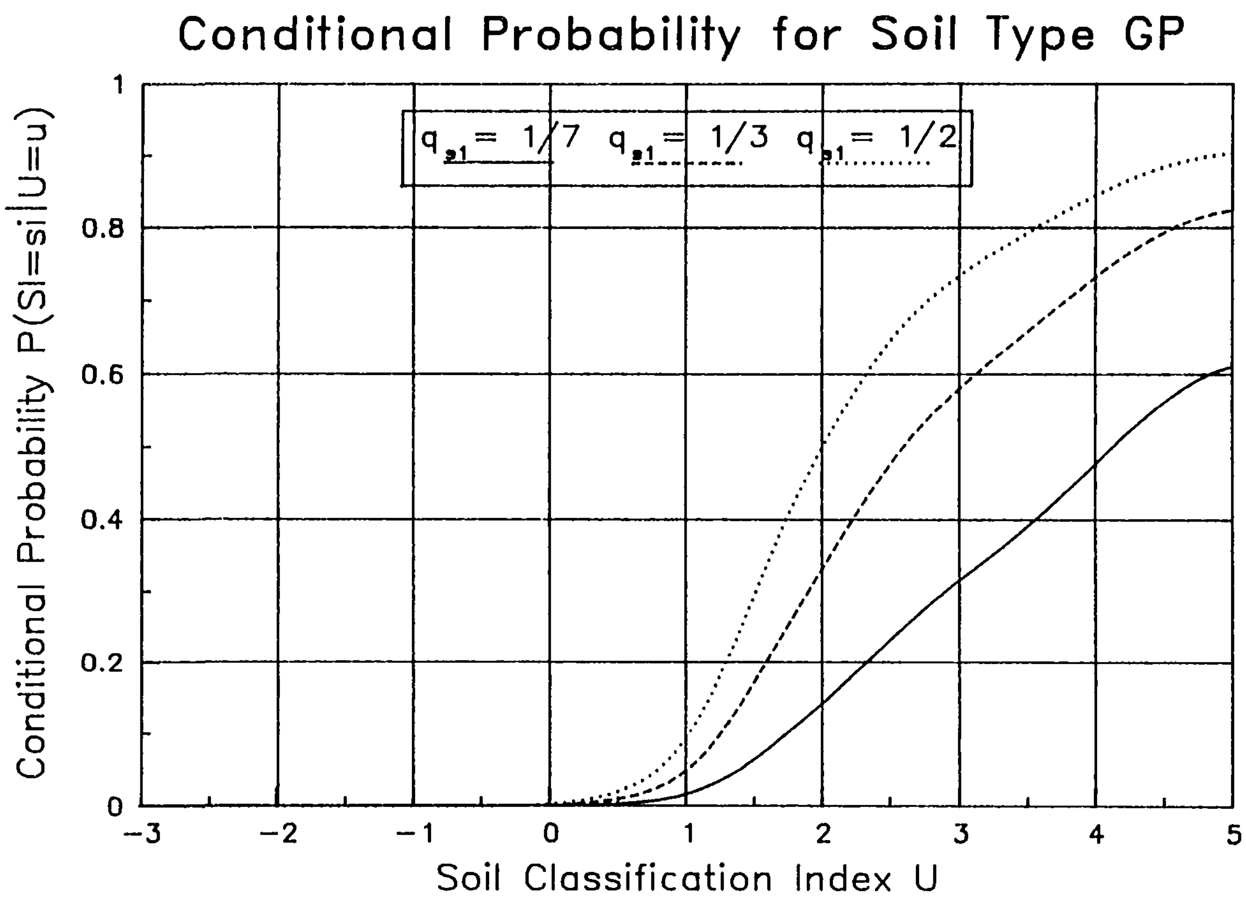


Figure 4.17 CPT Point Estimation Chart for Soil Type GP.

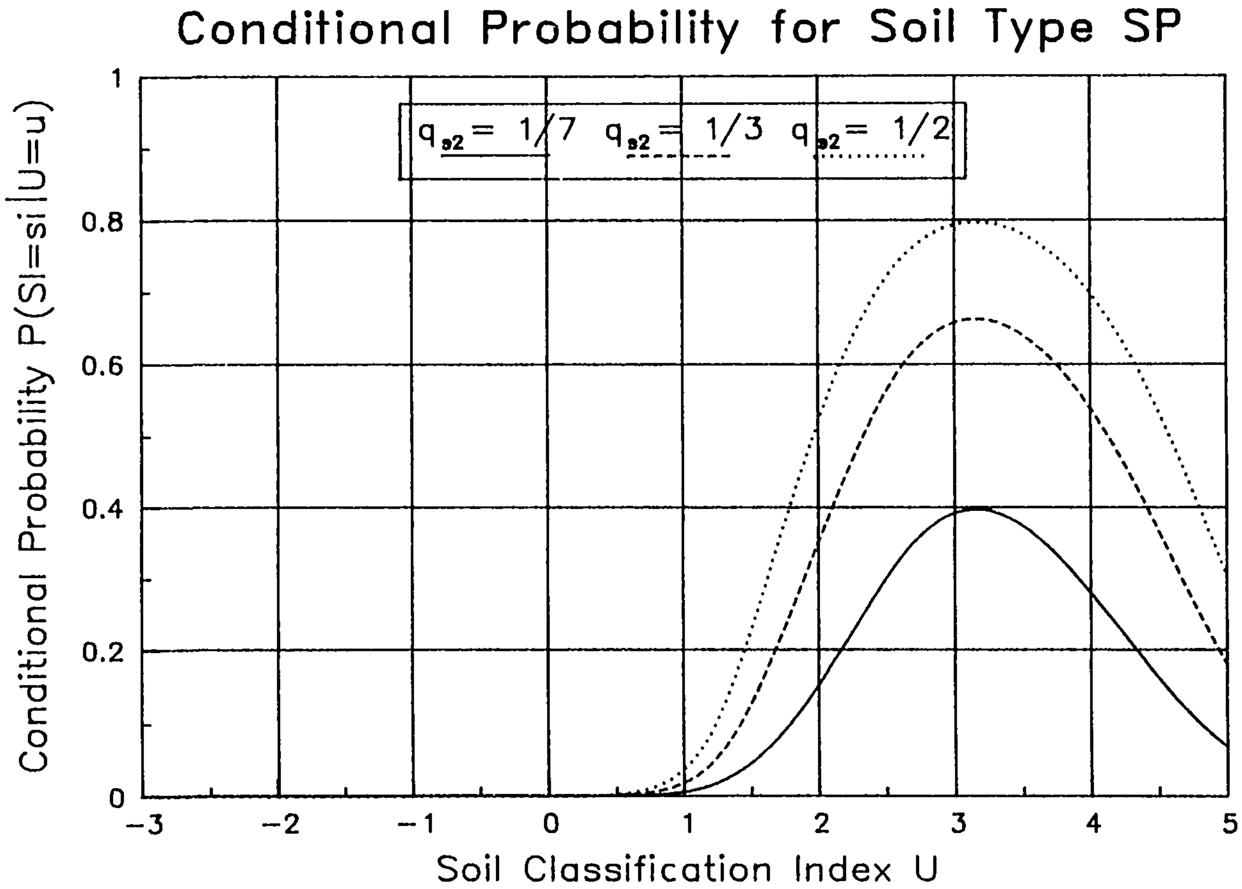


Figure 4.18 CPT Point Estimation Chart for Soil Type SP.

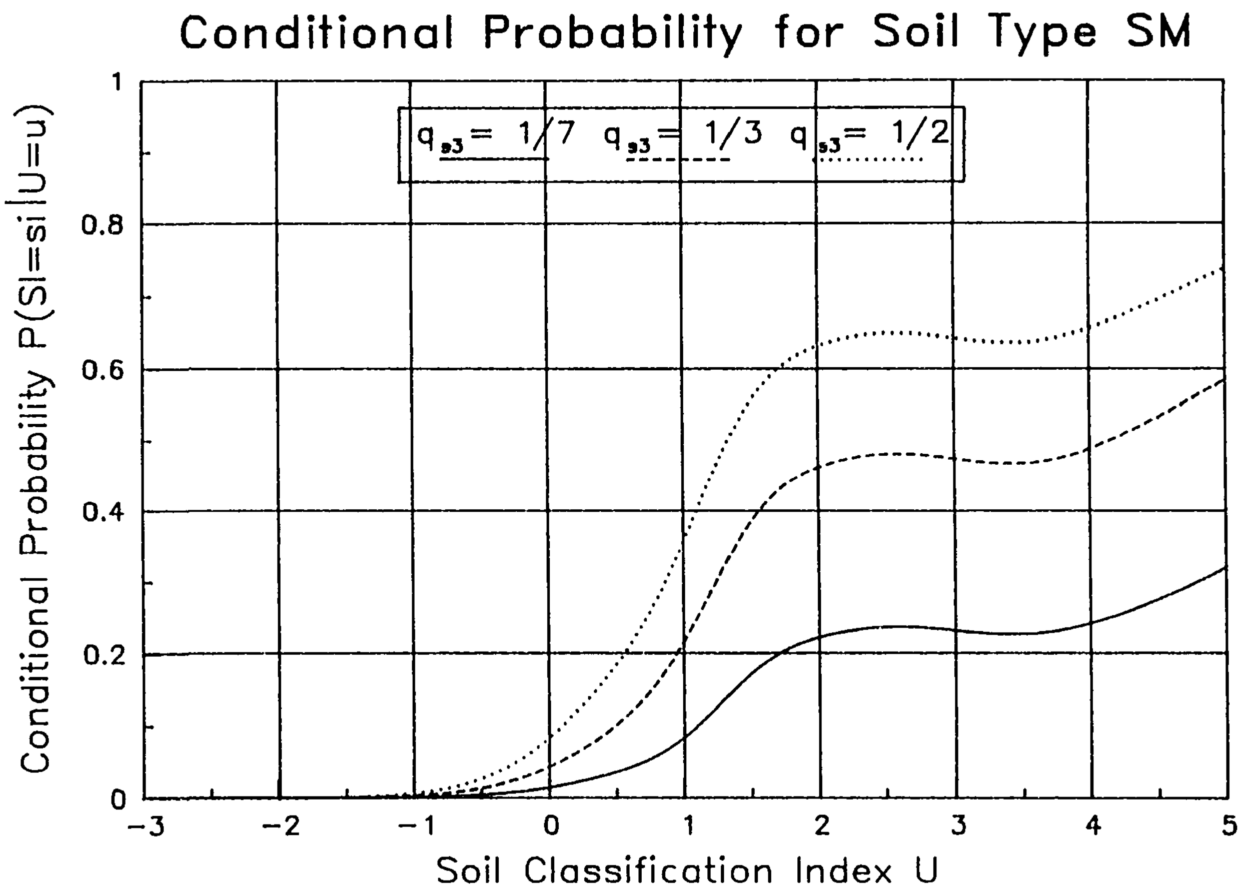


Figure 4.19 CPT Point Estimation Chart for Soil Type SM.

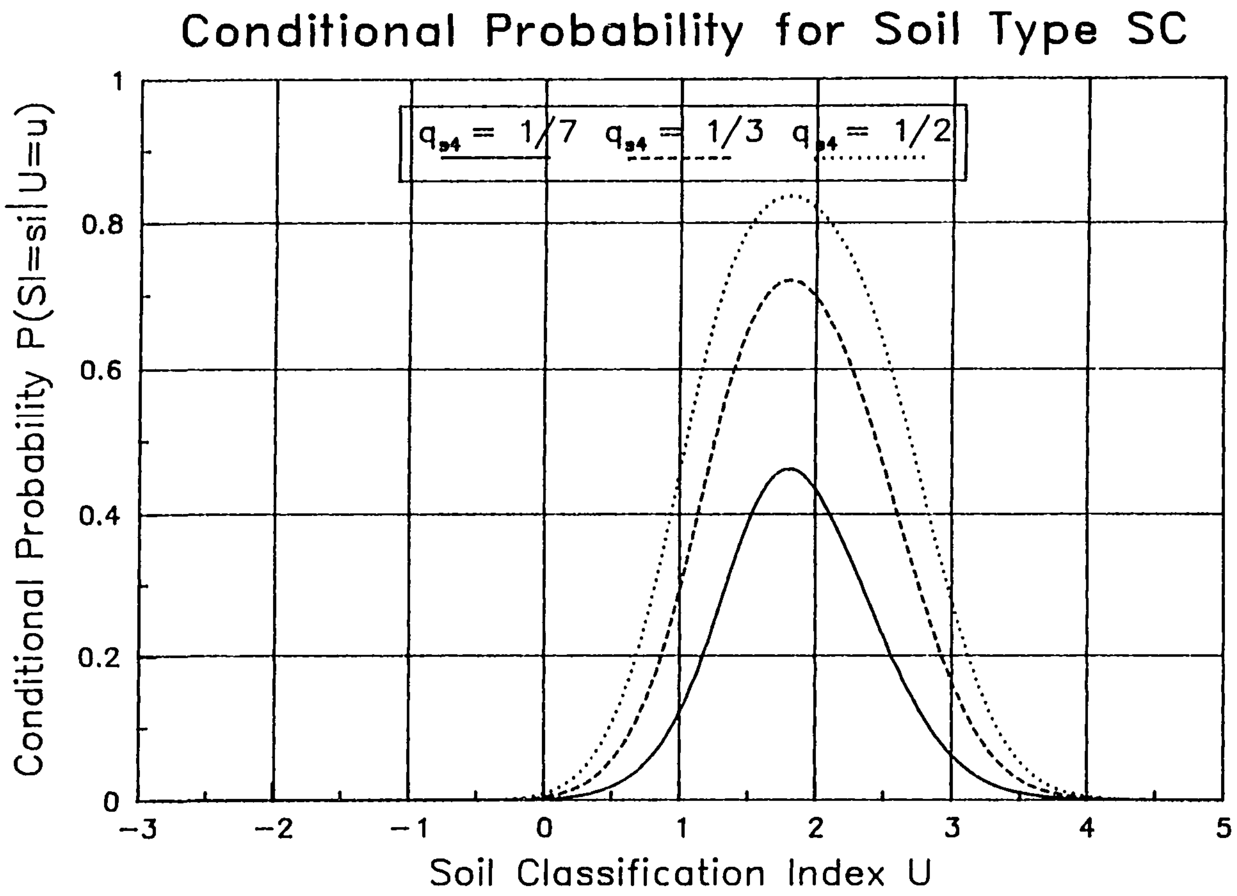


Figure 4.20 CPT Point Estimation Chart for Soil Type SC.

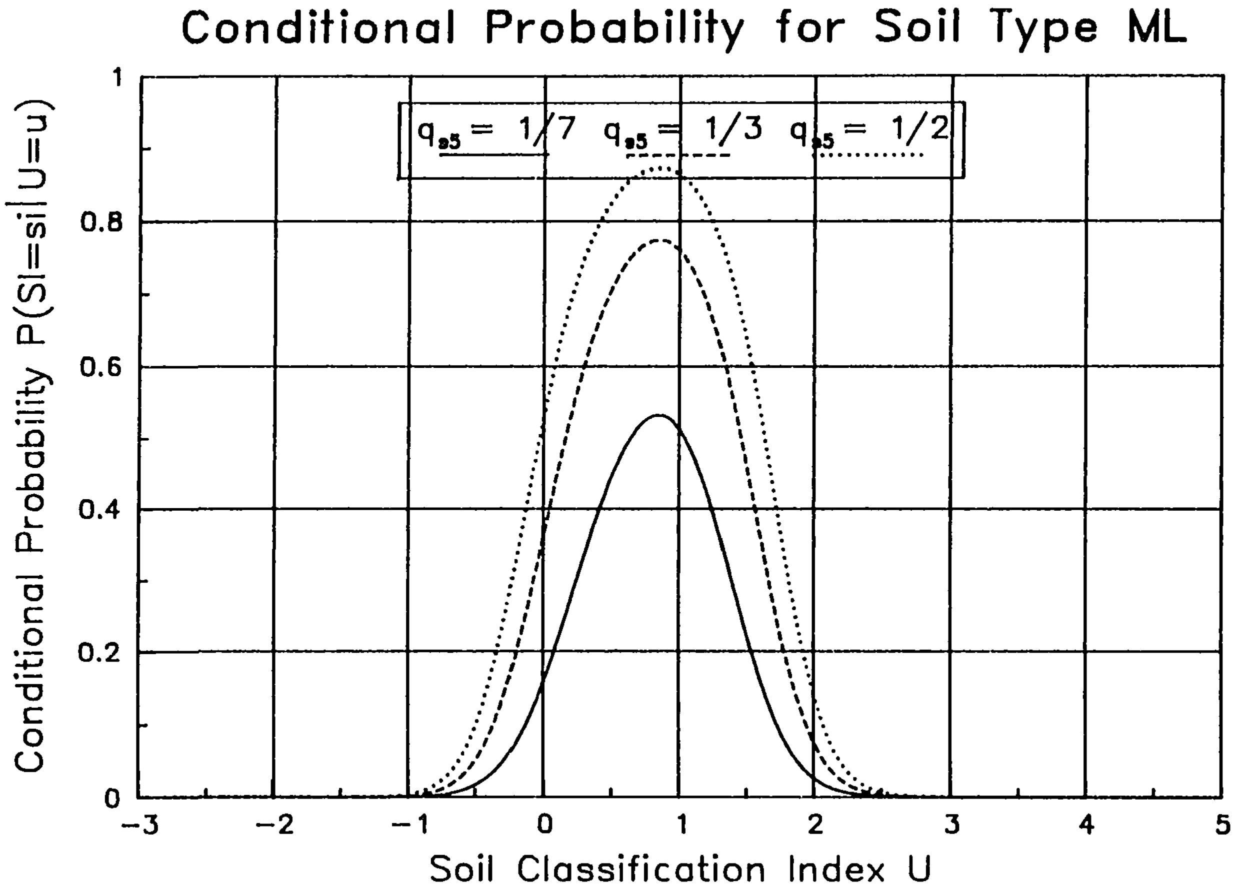


Figure 4.21 CPT Point Estimation Chart for Soil Type ML.

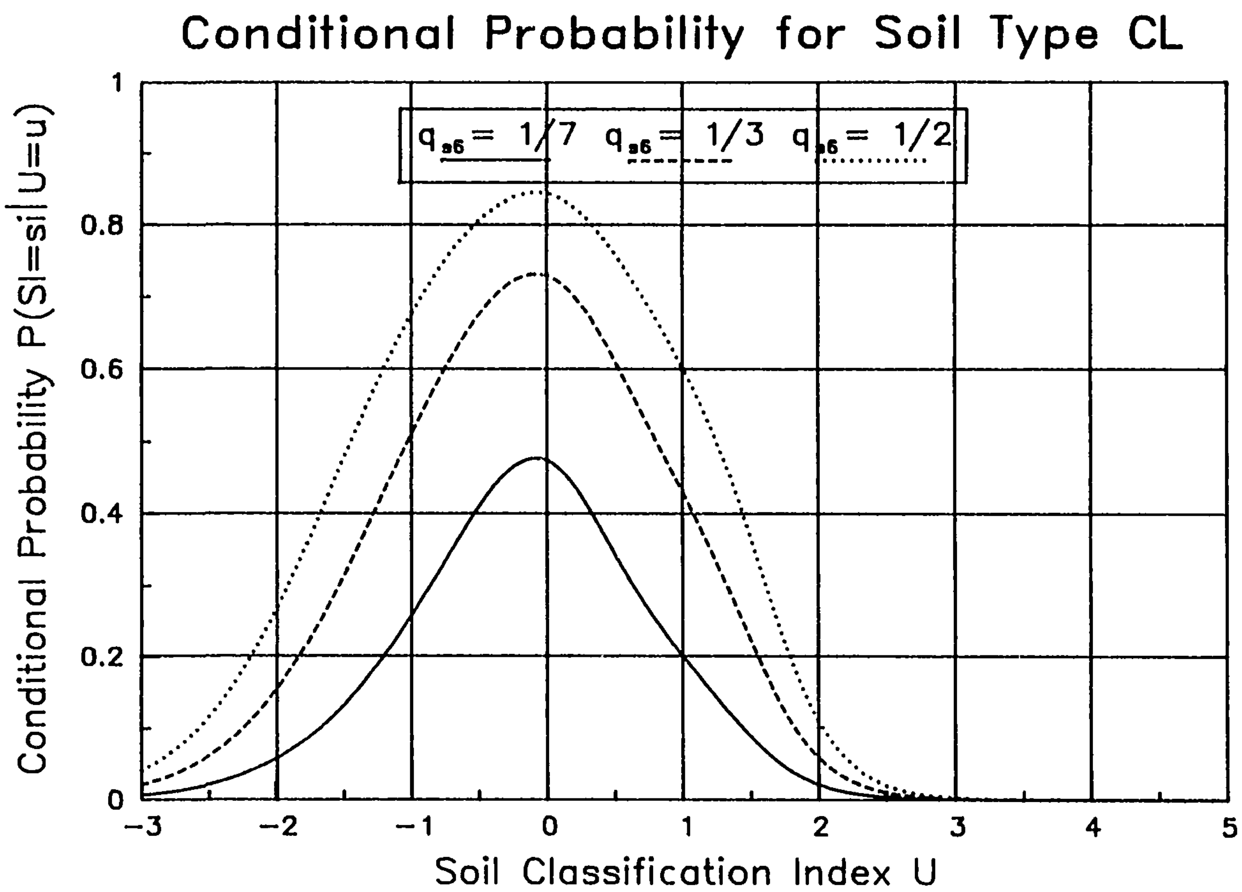


Figure 4.22 CPT Point Estimation Chart for Soil Type CL.

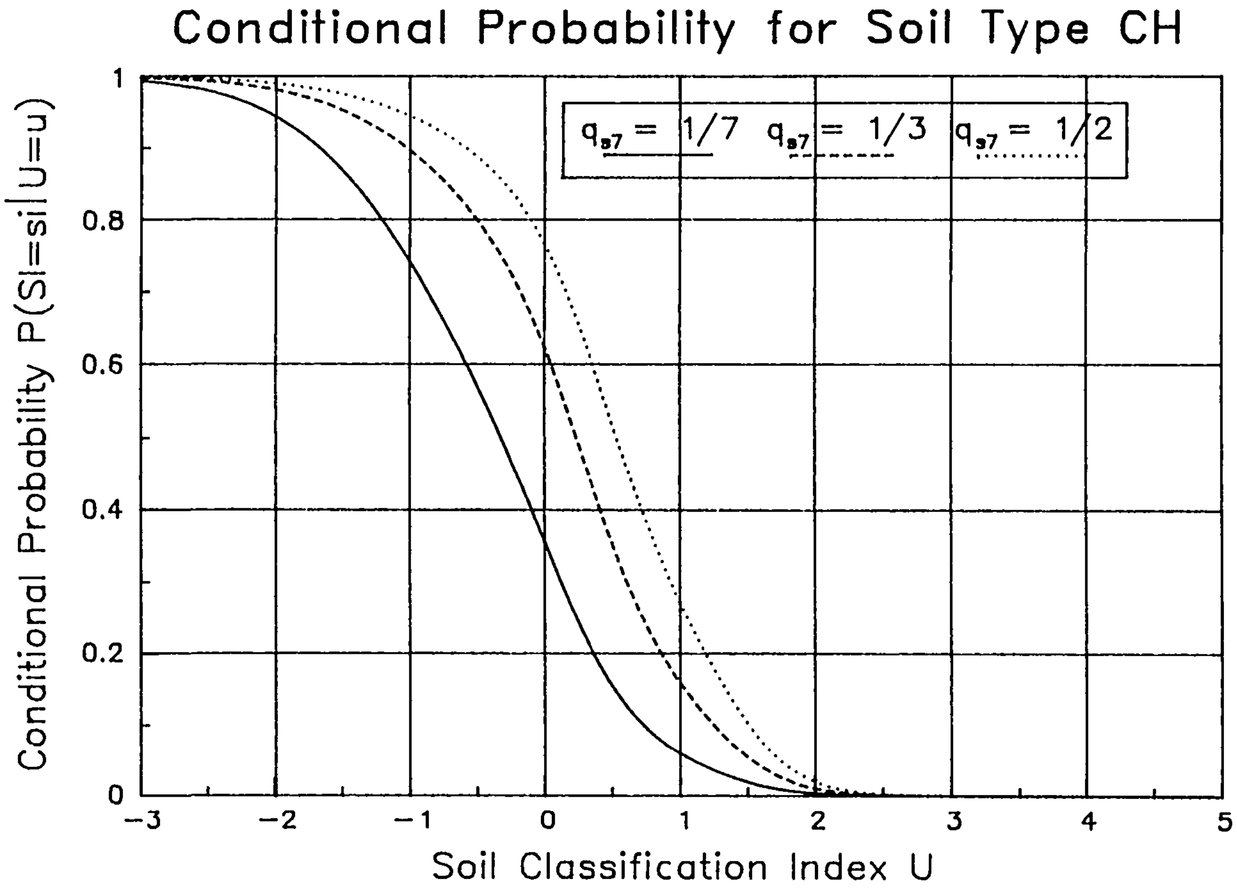


Figure 4.23 CPT Point Estimation Chart for Soil Type CH.

normal distribution well. More data are needed to check this result. Right now, the assumption of norma distribution is accepted with a confidence of 0.03. Then, Region Estimation and Point Estimation are developed accordingly.

Region Estimation is an approach similar to conventional soil classifications to classify soils. The predicted soil type of a soil sample depends on the region in which the soil behavior unit, U, of that sample falls. The classification criteria are shown in Table 4.2 and the corresponding probabilities are presented in Table 4.4 and 4.5. All the results are summarized in Figure 4.16. One important characteristic of this approach is that each region corresponds to several types of soils with different probabilities, not just one. It will let users make a decision of simplification. This is also true in the Point Estimation.

Point Estimation is an approach to classify soils directly by probability. It has an advantage of distinguishing different points within a region. In this sense, it will give a more accurate result than Region Estimation. However, extra information is needed before it can be used. This information is the probabilities q(si) (si = 1, 2, ..., M; M is the number of soil types in the concerned soil classification) with which different soil types are encountered in site investigations. Due to the restraint of in-situ data available, only assumed values of q(si) are taken here. The analysis results of Point Estimation are presented from Figure 4.17 to Figure 4.23. According to these charts, a point estimation of soil types can be performed if a **CPT** soil behavior unit, U, is given.

Finally, the significance of Region Estimation and Point Estimation can be summarized as follows. These two approaches indicate that when **CPT** data are used to identify soil types, the results can not be one hundred percent (100%) correct. Some disparity is possible. Therefore, if the problem at hand really concerns about soil composition and the classification results also indicate a possibility of unfavorable soil types, a boring test should be arranged to get the exact knowledge of the situation in-situ.

CHAPTER 5

FUZZY CPT SOIL CLASSIFICATION

5.1. Introduction

It has been suggested that uncertainty in general can be characterized in three ways: randomness, vagueness, and ignorance. Lack of causal repeatability of phenomena or of values of state variables, i.e., lack of the law of causation, is associated with randomness. The difficulty of defining concepts or state descriptions precisely, i.e., lack of the law of exclude middle will be considered as fuzziness. "If experiments are not repeatable, it is difficult for a subject to develop. Hence in the social sciences it is not easy to construct dependable theories. The extreme example is history where only very indirect experiments are possible. Thus increasing lack of repeatability of experiments could be associated with sufficient but not necessary conditions for lack of knowledge or ignorance" (**D.I. Blockley, 1985**).

In a **CPT** soil classification, only randomness and fuzziness are the relevant uncertainties. The randomness uncertainty has been discussed in Chapter 3 and 4. This chapter will begin with studying the problem of fuzziness. In this process, the fuzziness in a soil classification is discussed and the basic concepts and characteristics of fuzzy subset theory are reviewed first. Then the characteristics of a **CPT** soil classification are addressed and the possibility and necessity to apply the fuzzy subset theory to the classification problem are explored and examined. Also, as a first attempt, a temporary fuzzy **CPT** soil engineering classification is proposed. This fuzzy classification tries to solve not only the uncertainty of fuzziness but also, from another direction, the uncertainty of randomness in a **CPT** soil classification problem. Finally, a new package of **CPT** soil engineering classification based upon cone tip resistance q_c and friction ratio FR is suggested as a summary of the whole dissertation.

5.2. Fuzziness in Soil Classification

Fuzziness, as mentioned before, arises from a difficulty to precisely conceptualize natural objects, events, or states etc. Concepts are the abstractions of the natural existence. Any concept consists of two characteristic factors: connotation and extension. Because of its abstract nature, sometimes the connotation of a concept can be understood quite well. But the extension of it will look ambiguous when the concept is used to depict the corresponding things, phenomena etc in the natural world. This is the case for soil classification.

Although it is only in this century that soil mechanics has become a scientific discipline, human being's knowledge about soils has a long history. Based on the experience from a long term of practice with soils, the natural soils have been conceptualized into certain general groups, such as clayey soil and sandy soil. As the results of practical needs, detailed sub-division concepts of soil types have also been developed within each fundamental group, such as the ones described in **USCS**. All these concepts comprehend the knowledge of not only soil texture (grain size distribution) and physical characteristics, but also engineering properties (**Casagrande, 1948**).

It has been known that soils can have so many aspects of properties that it is not always easy for them to be classified properly into a group of comprehensive soil types. Actually, each aspect of soils has its own particular criterion to classify and a balance among them will depend upon the concerns (or priorities) of a classifier. Usually, different concerns will result in different boundary lines among soil types in most cases. Also remember that all the concerns (or priorities) generally have different engineering meanings and background. Therefore, to achieve a balance among them will not be an easy task. Sometimes it is very difficult to be accomplished. From this point of view, it can be said that more comprehensive the concepts of soil types are, more difficult a classification of soils will be. This situation will be reflected as a kind of uncertainty: fuzziness.

Another reason for the fuzziness in soil classification is that the compositional elements, physical characteristics, and corresponding engineering properties change gradually from one type of soil to another. It reflects the general rule of **"the changes of quantity will cause the changes of quality eventually"** in the natural world. In this sense, there are no clearly distinctive boundary lines among different types of soils. Soils can be sorted into several different groups only from a whole perspective of views. The difference between any two adjacent types of soils can be significant only when the "distance" between them, in certain scale, becomes large enough. In such a case, the fuzziness will also occur as the specific boundary lines among different soil types are portable.

Furthermore, not only will these two types of vagueness coincide but they will also mix with the uncertainty of randomness in a real situation of soil classification problem. It will make the problem more complicated. This may be one of the reasons why there were so many soil classification methods for almost the same engineering purpose, and why soil classification was the most confused chapter in soil mechanics. Therefore, it used to be, and still will be a challenge to develop a new kind of soil classification system.

Human beings have the capability to handle this challenge, just like L. A. Zadeh's observation. Humans can understand and analyze imprecise concepts which are not properly understood or emulated by existing analytical methods (**L. A. Zadeh, 1965**). Experts are able to handle all the uncertainties reasonably and intelligently because of their experience and genius. There is no vagueness in their minds. They make good balances among various aspects of soil characteristics and properties under different conditions. They can make correct decisions according to their criteria since they have enough experience with soils and a strong understanding of soils. The uncertainty only means to the not well trained mind. This may be one of the reasons why there are some uncertainty and flexibility in the original Airfield Classification System, which is the prototype of **USCS**, but the author was confident about it.

However, experts have difficulty expressing their certainty with the normal ways, i.e., the ways "supported by prevailing schools of epistemological thought that required precision as a sine qua non imposition on properly defined concepts". The formation of a classic (conventional) classification methodology is a rigid frame which does not allow a modification (adjustment) of a soil classification when conditions (concerns or priorities) change. But experts' certainty is based upon this kind of modification. Consequently, current soil classification systems can only transfer the certainty in an implicit way. As a result, it is not easy for other people later to appreciate this certainty in a definite way. This will cause the fuzziness in using these soil classification systems later on.

Furthermore, "the current methodologies showed a concern for precise representation of certain system aspects that were not only irrelevant to the analysis goals but that were an actual impairment in reaching the system understanding objectives". The classic (or conventional) classification methodology will give a precise soil classification on the surface since it complies with the law of exclude middle. However, such a precision will truncate the information from natural soil samples and produce some kind of distortional image (or simplified picture) on them. This point can be illustrated with the help of Figure 5.1. Although soil 1 and soil 2 are classified as the same type of soil with the same name, their similarity is not as good as indicated by the name of the soil type. The deviation between them can be quite large, depending on the individual

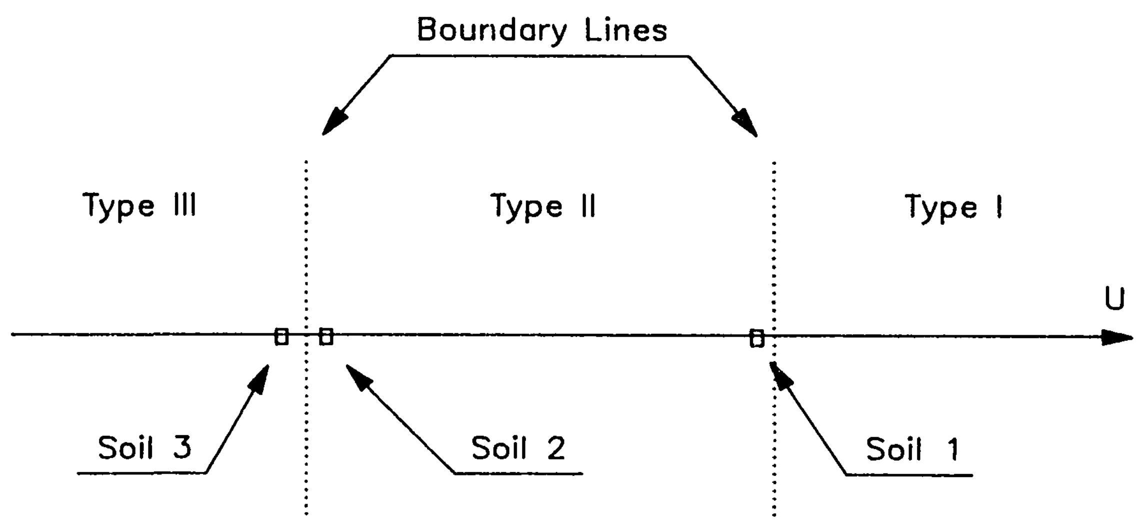


Figure 5.1 Exhibit of Information Truncation on a Crisp Soil Classification.

situation. On the other hand, it is not very easy to distinguish the difference between soil 2 and soil 3, but they have been classified into different types of soils.

Here, the problem is not the approach of truncation itself but the conditions in which to do it. It is clear that the validity of such a truncation caused by a classic classification methodology should be conditional in general. However, these conditions have no way to be reflected in current crisp classification systems. Therefore, such a conditional truncation of information from soil samples has been expressed as an unconditional one due to the inherent defect of the classic classification methodology. It can be imagined that the resulting distorted image of soils can later cause some misinterpretation of their properties.

Although the seriousness of the distortional image can be reduced to some extent by introducing more sub-types of soils, theoretically this problem can never be eliminated. It is embedded in the existing classic classification methodology because it always gives a "yes" or "no" answer. Also, more the soil types adopted, more complicated the soil classification, more inconvenient for use, and more uncertainty introduced.

The foregoing analysis of the fuzziness in soil classification has separated into two types: one is from **the comprehensive nature of a concept**, the other is due to **the law of "quantity change to quality change"**. They are actually two important characteristics in a soil classification problem. These two characteristics will not cause any problem if we can handle them properly. However, what we have done is quite the opposite. We looked at them in view of the conventional methodology. We used this methodology to describe them in order to get a "precise" representation of a comprehensive soil classification system. This practice has put us in a dilemma called fuzziness or vagueness. This is because the conventional crisp methodology adopted in current soil classifications has impeded a clear transference of the relevant information and knowledge which has been used during the system development. Thus, the interpretation of this vagueness will depend on personal experience and training, "so that it may be said that there are as many classification systems as there are engineers using them" (**Waterways Experiment Station, vicksburg, Mississippi, 1951**). All these phenomena have indicated the objective basis to use a new methodology which already exists in the state of art: fuzzy subset theory.

5.3. Basic Concepts and Properties of Fuzzy Subset Theory

The basic notion of fuzzy subset was first introduced by L. A. Zadeh in 1965. It was the result of his consideration of the big difference between human brain's capability and existing analytical and modelling methods, and of the contradiction between precision and correctness in an analytical modelling. As Zadeh pointed out latter, "the theory of fuzzy subset is, in effect, a step towards a rapprochement between the precision of classical mathematics and the pervasive imprecision of the real world - a rapprochement born of the incessant human quest for a better understanding of material processes and cognition" (**A. Kaufmann, 1975**).

The theory of fuzzy subset is based upon the concept of degree of membership u(x), also named the degree of belongingness or characteristic function. This concept was generalized from the classic notion of a set. As a fundamental concept in mathematics, a set A is a collection of elements in a considered domain. Any element x in this domain can only either belong or not belong to A, i.e., the law of exclude middle is observed. In mathematics, this fact can be described by a membership function or characteristic function $u_A(x)$, whose value indicates (yes or no) whether x is a member of A:

$$\mu_A(x) = \begin{cases} 1.0 & if\ x \in A \\ 0.0 & if\ x \notin A \end{cases} \tag{5.1}$$

Here, the symbol $\in$ means "belong to".

Zadeh believed that "much perhaps most, of human cognition and interaction with the outside world involves constructs which are not sets in the classical sense, but rather 'fuzzy sets' (or subsets), that is, classes with unsharp boundaries in which the transition from membership to nonmembership is gradual rather than abrupt" (**A. Kaufmann, 1973**). This idea can be described mathematically as $u_B(x) \in [0, 1]$, i.e. $u_B(x)$ can be any value between 0 and 1. This range of values represents that an element may belong to a set in any degree from non-membership to partial membership and to full membership. Therefore, a fuzzy subset can be defined mathematically by the expression (**A. Kaufmann, 1973**)

$$B = \{ x \mid \mu_B(x) \}, \qquad \forall x \in A \tag{5.2}$$

Here, A is a classic set, enumerable or not, and x is an element of A. $u_B(x)$, called fuzzy membership function, is the grade or degree of membership with which x is belonging to B. On the other hand, a classic subset B_α can further be elaborated based upon a fuzzy subset with level α, here $\alpha \in [0,1]$. This classic subset can be determined as

$$B_\alpha = \{x \mid \mu_B(x) \geq \alpha \} \tag{5.3}$$

In this way, a fuzzy subset can be easily "transferred" to a crisp subset which has sharp boundary lines.

It is apparent that the most important thing to find out for a fuzzy subset is a proper fuzzy membership function for it. This kind of membership function is supposed to express human's comprehensive cognition and quantification of the gradual change of certain abstract objects in a concerned domain. Since this abstract comprehensive quantity can not be measured directly, the membership function of a fuzzy subset has to be estimated. This estimation will be affected by personal experience, training, the capability of quantification and other factors, such as psychology, responsibility, etc (**Brown et al, 1985; Santamarina, 1987**). Thus, a membership function seems uncertain in form and every individual can have a different one. Since people always set out from prior knowledge to understand new ones, a required mathematical description on a gradual change will remind people the format of the probabilistic theory. This might be one of the reasons that the fuzzy subset theory is easily confused in concept with the probabilistic theory.

The fuzzy subset theory and probabilistic theory are totally different in nature. What a fuzzy subset wants to describe is certain in the sense of randomness controlled by the rules of probability. The problem of uncertainty described by fuzzy subsets stems from the way humans cognize the natural world. Humans are always trying, at least in the sense of knowledge, to make the natural world in order, which may be the one aspect of subjective initiative. They always try to simplify this world by using all of the methodologies he has learned from their experience in their process of cognizing the natural world. Among them are comprehension, abstraction, and quantification in subjective. All these approaches together can produce very comprehensive, abstractive, and flexible quantities that can not be well expressed in a usual or classic way but can be felt and so are fuzzy.

Certain characteristics can be expected in the process of generation and evolution of a fuzzy quantity. First, this process is a series of iterative cognition cycles which never finish at once and also theoretically never end. The results from a foregoing cognitive process will be checked and tested in the following practice and the feedback will be used to modify the subjective cognition. As a simulation of a fuzzy quantity, the development of a membership function should also follow this iterative patten due to the nature of estimation and evolution of the fuzzy quantity.

Second, a subjective quantification in a cognitive process has a characteristic of flexibility which is only restrained by the purpose of this process. In other words, whatever the results will be, the only requirement is to serve our purpose. This is a general characteristic of subjective activities and is contradictory to the basic philosophy of classic mathematics. In this sense, it is not easy to find a proper form for the quantification in classic mathematical language humans have used to describe the laws of nature, specifically the form of a membership function for a fuzzy subset. As another side of the flexibility, a fuzzy subset in a fuzzy approach may have several membership functions with different forms but all good for using. This can be proven by the fact that different people may think in different patterns but they all can be successful engineers. From this point of view, it is questionable whether the proposition of a probabilistic theory based upon fuzzy sets can provide a correct and useful tool to solve mixed uncertainty problems.

Conclusively, due to the impossibility of directly measuring human's quantification in a process of cognition, the determination of a fuzzy membership function should be a procedure of "trial and error". In the beginning, this kind of function can be established with some "rule of thumb" estimation. It is obvious that the quantification in human's cognition process is one thing and expressing this quantification in mathematical language is another. Therefore, being an experienced and successful engineer might not guarantee a good membership function for a kind of vagueness in which he or she has an expertise. Also averaging different persons' membership functions does not secure a good one since this method only makes sense in statistical theory. No matter how a membership function of a fuzzy subset is obtained, the correctness of it has to be checked in practice. A modification of it can be made, if necessary, based on the feedback. This procedure should be considered as a general rule to determine a membership function.

It is believed that the fuzzy subset theory has supplied a tool to summarize human's experience to deal with fuzziness. However, the experience obtained so far on this theory still can not guarantee a proper use of it in a new field. It may be appropriate to say that the significance of the fuzzy subset theory lies in that it has provided an access to find out an unified, precise, mathematical language to describe fuzziness. Based on this language, a proper simulation of fuzziness can be worked out through an iterative process between theory and practice.

5.4. Basic Concepts of a Fuzzy Soil Classification

Fuzzy subset theory has many practical uses. One of them is to describe a fuzzy concept, the extension of which is not very clear. The physical meaning of a fuzzy subset is defined by a connotation of the fuzzy concept. The extension of the fuzzy concept is then described by the corresponding membership function of the fuzzy subset. This is the way the fuzzy subset theory is used in a soil classification. In such a system, some core soil types, such as the ones in USCS, are decided first. Each type of soil is then described by a corresponding fuzzy membership function and certain manipulations of these fuzzy membership functions are performed according to some rules of the fuzzy subset theory. Also, any other kinds of concepts about soil types in fuzzy sense, such as liquefiable soils, frost-susceptible soil, etc, can be introduced as the supplementation to the core soil types of the fuzzy soil classification. Their fuzzy membership functions can be worked out similarly to the ones of the core concepts. Because of a possible independence between the connotations of the adopted core concepts and other supplemental concepts, their membership functions do not need to have direct relations.

The advantage of using fuzzy methodology in a soil classification problem is the unity of content and format presented in such a system. As we already know, soil types are all fuzzy concepts but the formation used to express them is crisp within current conventional soil classification systems. This problem is now eliminated by using the fuzzy expression, as the gradual changes of soil types can be described explicitly by corresponding fuzzy membership functions.

Also, since fuzzy subsets can be transferred into classic subsets if a group of α level values are given, all conditional truncations on soil type information can be described explicitly now. In this sense, it can be said that a fuzzy soil classification system is the most general form of a soil classification for a fixed set of soil classification indices. The differences among different concerns or priorities of soil classification are included in and reflected by the values of level α. From this point of view, a flexible nature (or uncertainty) of current soil classifications is exposed again. Sometimes, it is not easy to find a convincible reason to chose certain values of level α for an ordinary crisp soil classification.

One more thing which is worth being mentioned, is that this kind of transfer relationship between a fuzzy set and the corresponding classic subsets is desirable to simulate a phenomenon of "concept shifting" which occurs during people's thinking. When people are talking about soil types in general sense, the concepts of various soil types are fuzzy. However, when the discussion is actual-problem-oriented, the concepts of soil types may be crisp in nature. In these two cases, people use the same terms but the meanings of them are slightly different.

Furthermore, the physical meanings of α values can be any classification criteria which people need to serve certain purposes. For example, with the culmination of human's experience and knowledge about soil engineering properties, some particular criteria for certain engineering problems can be established by experts. These problems could include strength stability problems, settlement stability problems, liquefaction, compaction, seepage problems, etc. All the criteria for them are represented by certain values of α. In this way, experts' experience and knowledge about soils can be recorded in a succinct and explicit style. These experience and knowledge later can help other engineers to make preliminary decisions during the initial stage of a problem solving process. In such a case, each of these criteria is supposed to give only an indication from a certain point of view for possible soil properties because of their possible complexities. If necessary, some follow-up tests can be performed later to know the exact soil behaviors in an in-situ situation. In this way, the fuzzy methodology also supplies a general basis for developing some kinds of subjective empirical relations between engineering properties and soil types, which is the summation of the expertise currently available.

It is understandable, however, that searching for reasonable criteria mentioned before will be a long accumulative procedure due to their empirical nature. Therefore, this topic should be concentrated on in future research. From now on, every person can give his or her contribution to solve a particular problem by addressing his or her opinion on an interested criterion. After a certain period of time, a general discussion of it will be very helpful to come to a tentative agreement. The criterion will then be checked in practical usage and the feedback will be used to work out its new version for the future.

In general, a fuzzy soil classification has some special advantages and properties over current conventional soil classifications. It will provide a scientific methodology to summarize and express human's experience with soil types. It is worthy to develop such a classification although only a tentative draught of it can be suggested here because of its iterative nature.

5.5. Implementation of a CPT Fuzzy Soil Classification

It has been pointed out in Section 5.2 that there is some fuzziness in soil classification due to **the comprehensive nature of soil type concepts** and **the law of quantity change to quality change**. The conventional crisp methodology has no capability to describe and reflect such fuzziness. However, people will not spend money and time on current soil classification systems just for solving this covered dilemma, which is understandable because engineers have already been accustomed to existing soil classification systems. Also, this problem can be lessened or even eliminated by personal judgements of well trained and experienced engineers. Therefore, those conventional soil classifications of compositional type will not be changed in the future. With such a prediction, the discussion of the fuzziness in soil classification here will in general help us establish a flexible attitude on those conventional systems and properly use them.

The situation will be different when a new soil engineering classification is required to be worked out. In such a case, new techniques should be, whenever possible, integrated and merged with conventional ones so that a brand of new classification system can be worked out. Otherwise, no progression of technology can be expected.

Concerning our specific case, it has been pointed out previously that the fuzzy subset theory is a scientific way to describe and express a soil classification system. This theory should be adopted when a soil classification system of cone technology is developed. However, implementation of this general idea in our case is not as easy and clear as the theory itself appears. The fundamental problems to solve are what core concepts should be adopted and how the corresponding membership functions are determined. Before finding solutions for these problems, it will be helpful to discuss following characteristics of our **CPT** soil classification problem.

As we have illustrated before, a **CPT** soil engineering classification can be affected by many factors. However, it seems that all of these factors have practically no way to be distinguished and evaluated separately and independently based solely upon cone measurements, even in a subjective sense. All their effects are mixed and can only be seen together in the results of cone measurements, i.e., q_c and FR (from f_s) in a **CPT** case. If no other types of experiments are performed, no other independent source of information of soil types will be available. Therefore, our soil classification in a **CPT** case is basically a two-dimensional problem (q_c and FR). This is the first characteristic we would like to point out.

Second, based upon the assumptions and approaches suggested in Chapter 1, this two-dimensional problem can be simplified into an one-dimensional one. In this one-dimensional problem, the parameter used is the soil classification index, U, defined by Equation (2.1), (2.2), and (2.3). Of course, the degree of correctness of each assumption in this simplification will be different. However, due to the author's limitation of experience with this matter, the influence of these assumptions have no way to be estimated quantitatively for the time being, neither objectively nor subjectively. As a result, the problem at hand is primarily treated as an one-dimensional classification problem.

The third characteristic of the problem can be identified by checking the procedures with which current **CPT** soil engineering classification charts were developed. As we have already discussed in Chapter 3, existing **CPT** soil classification charts use soil behavior (soil responses to cone penetration) measurements as soil classification parameters. However, the classification criteria adopted are transferred from compositional soil classifications, such as the **USCS**. As a natural result of this practice, the same naming system has been employed, too. This approach to elaborate a soil classification has, as mentioned in Chapter 3, the advantage of keeping the continuity and integrity of the whole soil classification system. Therefore, the accumulated past experience will be related to the new classification to develop. Nonetheless, it has also brought the explicit uncertainty analyzed in Chapter 3 and 4 into the new system. This uncertainty has shown itself as the overlaps among different soil types in **CPT** soil classification charts.

It is obvious that due to those overlaps among different soil types, soil types identified by a **CPT** soil classification are not exactly equivalent to the corresponding ones identified by a compositional soil classification although they have same name. A soil type identified by a classification of behavior type corresponds to several different compositional types of soils. So, as a type of soil, it has a property which the corresponding soil type identified by a compositional classification does not have. No matter whether people realize or not, it is a fact. The name sharing has created such a problem. The information that there are some overlaps among different soil types in soil classifications of behavior type has been erased. Therefore, this practice will cause some confusion both in theoretical and practical points of view.

On the other side, even if the uncertainty in a **CPT** soil classification is handled by some statistical approaches, such as the ones suggested in Chapter 4, the problem is still not totally solved from a practical point view. For example, people will still feel uncomfortable when we say that a soil layer identified by **CPT** is sand with a probability of seventy percent (70 %) and silt with the rest thirty percent (30 %). When we do this, we actually put an emphasis on the uncertainty, soil composition. Notice that in a compositional soil classification, soil composition is certain, but the related behaviors can be predicted only in a general sense. There is uncertainty in anticipating a specific engineering property based solely upon a soil composition. People feel comfortable about that because the emphasis in that system is on the certainty: soil composition.

The same philosophy can be followed to develop a **CPT** soil classification. In such a system, an emphasis can be put on the certainty, soil behavior. Soil composition can be predicted according to behavior but with some uncertainty. A precise knowledge of it can only be obtained from a boring test. If we do this, there will be no uncertainty in a form of overlaps among different soil types in such a classification system. The uncertainty discussed previously can be transferred into a factor which can influence the choices of boundary lines among different types of soils in this new system. Keeping in mind the general purpose of a soil engineering classification, this practice should be acceptable.

The above consideration can be materialized by following two strategies to develop a soil classification of behavior type. The first one is to make a behavior type of system truly a behavior type. Not only the soil classification indices but also the classification criteria are truly based directly upon soil behaviors. This kind of classification will serve the cases where it really does not matter what the composition of soils is, as long as the knowledge of soil behaviors is available. The problem for this strategy is that the new system will be isolated with the old ones. Therefore, the usefulness of it will be limited substantially.

The second strategy is the result of a modification of first one. In this strategy, the classification criteria should directly include not only the information of soil behaviors but also their correlations with the corresponding classification criteria of compositional type. One thing that is common in both strategies is that a different naming system should be adopted. In the case observing second strategy, the new naming system should reflect the fact that one soil type identified by the new soil classification of behavior type corresponds to several soil types identified by an old compositional classification.

The second strategy will be observed in this research since it is more acceptable. This can be implemented by the fuzzy subset theory in a straight forward way. Each aspect (including engineering properties) of soils can be empirically correlated with the **CPT** soil classification index (soil behavior unit), U, and a corresponding fuzzy membership function can be worked out in some way. Then, some comprehensive soil types can be defined in a systematic way simply by following the rules of manipulations on the functions (fuzzy subsets).

In this dissertation, due to the limited time and data available, only a fundamental part of this new system is chosen and focused upon. This part is to determine the fuzzy membership functions indicating the correlations between soil types identified by classifications of compositional and behavior type. This can be accomplished by borrowing the probabilistic density functions of compositional soil types as the prototypes. Then the required membership functions are defined by normalization and modifications based upon currently obtained experience. The resulting membership functions together will describe the gradual variations of soil types, as exhibited in next section.

5.6. Frame of a CPT Fuzzy Soil Classification (CFSC)

The following fundamental (or core) fuzzy concepts of soil types are adopted temporarily as the outcome of considerations based upon previous analysis on a **CFSC**. They are High Probability Clayey Soils (**HPC**), High Probability Mixed Soils (**HPM**), and High Probability Sandy Soil (**HPS**), as shown as

SOILS

HPS	**HPM**	**HPC**
$\mu_s(U)$	$\mu_m(U)$	$\mu_c(U)$

Here $\mu_c(U)$, $\mu_m(U)$, and $\mu_s(U)$ are the fuzzy membership functions of **HPC**, **HPM**, and **HPS**, separately. They are estimated according to some up-to-date experience. Temporarily, two aspects have been taken into a account.

First, it is assumed that there is core center or range where the membership function will take a value of one (1) for each of these fuzzy subsets. The absolute change rate of the abstract fuzzy belongingness at any point is supposed to increase as the "distance" between the point and the center of the fuzzy subset increases up to certain limits.

Thus, a kind of S shape function, which satisfies this requirement, is employed for these fuzzy membership functions. Such results should approximately reveal the changes from quantity to quality. Second, the positions of the fuzzy membership functions for the three soil types are decided by a perspective consideration of overall soil properties. What the membership functions want to show is that soils in these three groups have fundamentally different engineering properties but with no sharp boundary lines among them. The changes are gradual from one type of soils to another.

Although the previous discussion does not specify a way to determine these three fuzzy membership functions, it has described what we should get. It actually plays a role to provide some directions. Empirically, the three membership functions can be determined based upon the data (relative frequency of U) given in Figure 3.8. In that chart, soils have be reorganized into three groups (Group I, Group II, and Group III), which are directly related to **HPC**, **HPM**, and **HPC** here. The density functions of the three soil groups are assumed to be normal and are

$$f_s(u) = \frac{1}{0.834586\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{u - 2.6575}{0.834586}\right)^2\right) \tag{5.4}$$

$$f_m(u) = \frac{1}{0.724307\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{u - 1.35}{0.724307}\right)^2\right) \tag{5.5}$$

$$f_c(u) = \frac{1}{0.86332\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{u + 0.1775}{0.86332}\right)^2\right) \tag{5.6}$$

The validation of this assumption is checked as similarly as before in Section 4.2. The corresponding results are shown in Figure 5.2, 5.3, 5.4, 5.5, 5.6, 5.7, and Table 5.1. They are given just for a reference purpose. The explanations on them can be figured out by referring the same discussions given in Section 4.2.

Table 5.1. Results of Distribution Fitting Tests of Combined Soil Types

Soil Type	Distri.	mean	S. D.	Chi-Square Test			K - S Test		
				Est.	D. F.	S. L.	Dplus	Dmi	S. L
HPS	f_s	2.66	0.83	21.46	8	6.0E-3	0.096	0.123	0.098
HPM	f_m	1.35	0.72	79.01	11	2.3E-12	0.154	0.096	0.017
HPC	f_c	-0.18	0.86	13.68	8	9.0E-2	0.074	0.119	0.118

Note:

Distri: Distribution Function;
S. D.: Standard Deviation;
Est.: Estimate;
D.F.: Degree of Freedom;
S.L.: Significance Level;
Dplus: The maximum positive deviation of the empirical cumulative distribution over the hypothesized cumulative distribution function;
Dmi: The maximum negative deviation of the empirical cumulative distribution over the hypothesized cumulative distribution function.

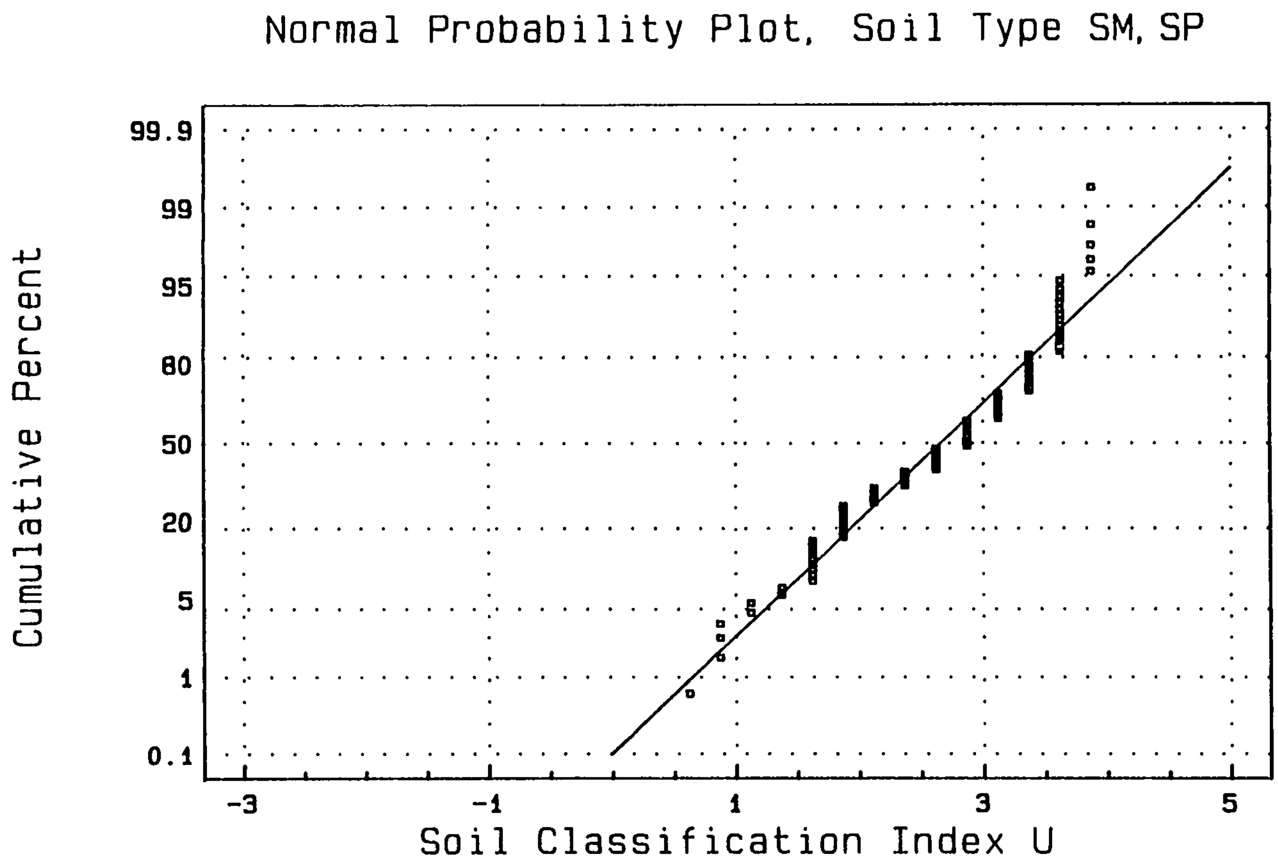


Figure 5.2 Normal Probability Plot for Soil Group SM, SP.

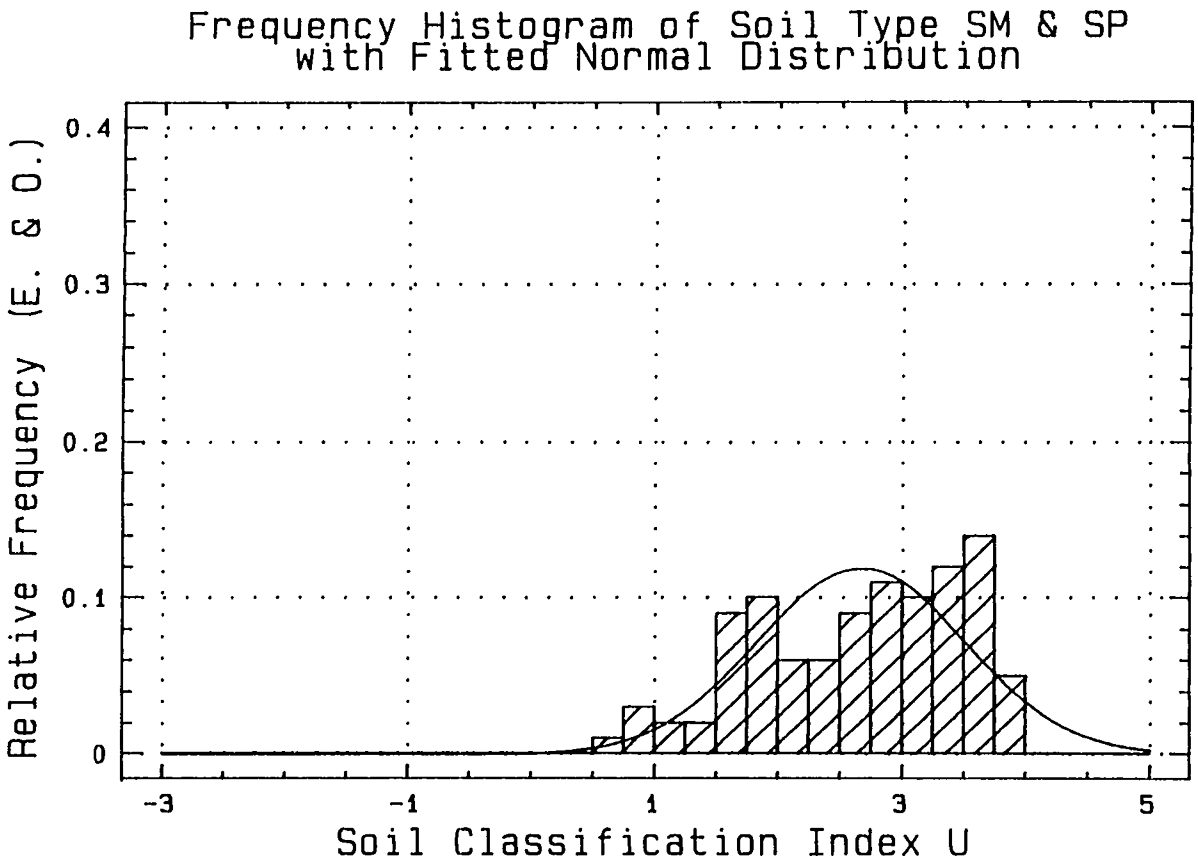


Figure 5.3 Frequency Histogram for Soil Group SM, SP.

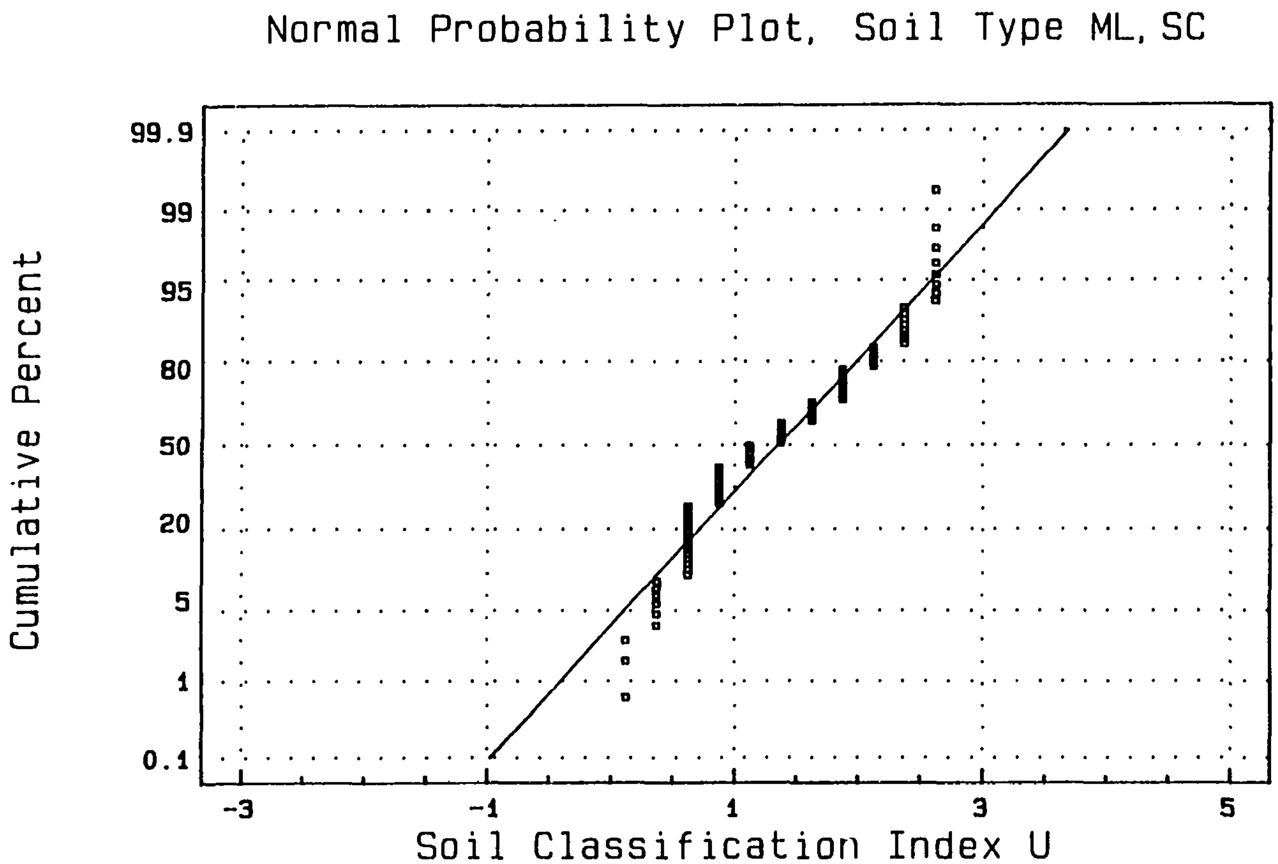


Figure 5.4 Normal Probability Plot for Soil Group ML, SC.

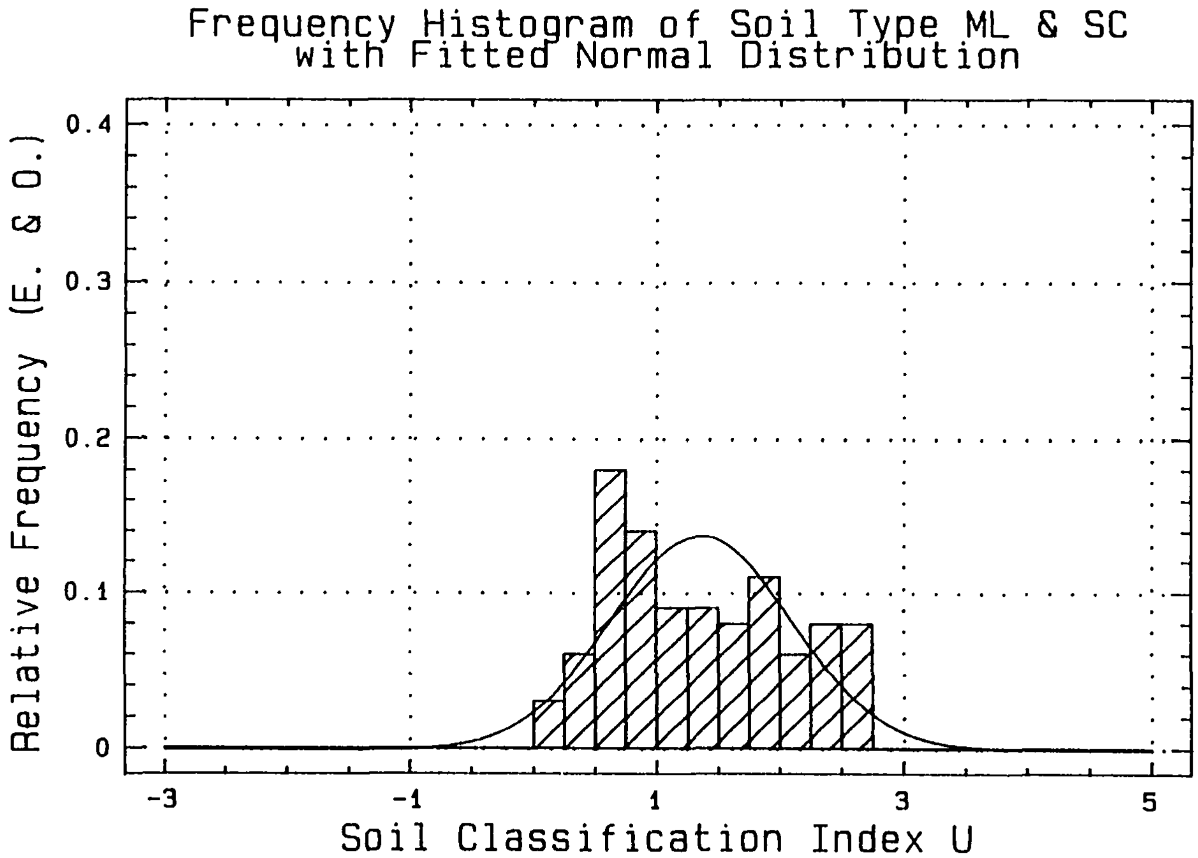


Figure 5.5 Frequency Histogram for Soil Group ML, SC.

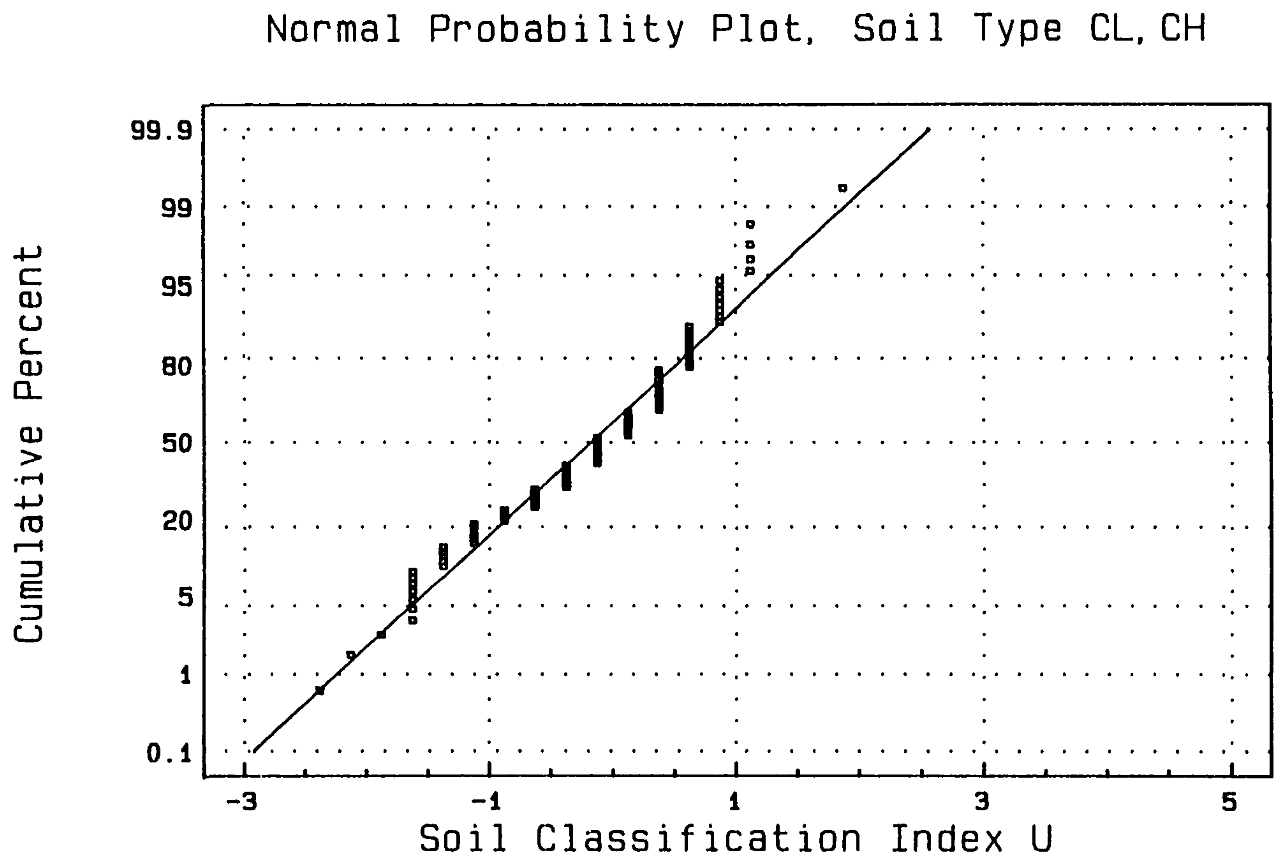


Figure 5.6 Normal Probability Plot for Soil Group CL, CH.

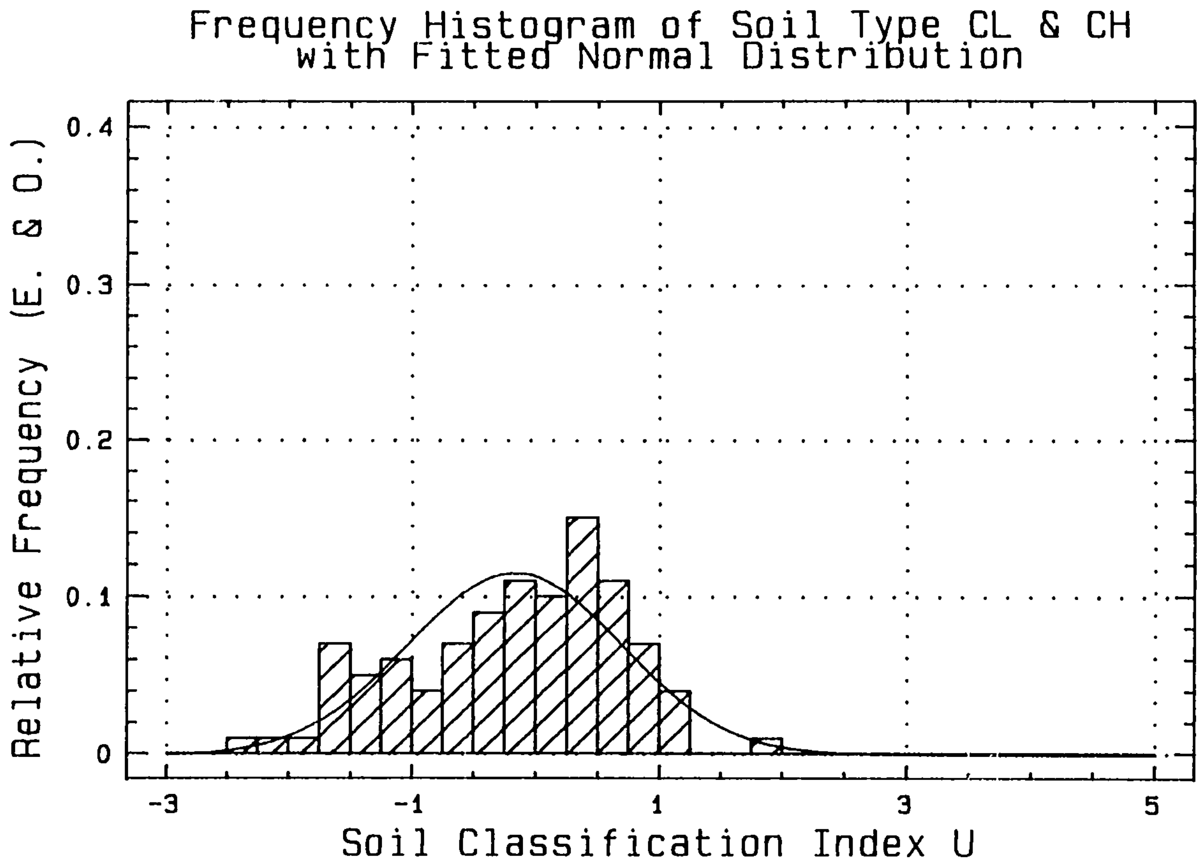


Figure 5.7 Frequency Histogram for Soil Group CL, CH.

Since

$$f_s(u)\,|_{max} = \frac{1}{0.834586\sqrt{2\pi}} \tag{5.7}$$

$$f_m(u)\,|_{max} = \frac{1}{0.724307\sqrt{2\pi}} \tag{5.8}$$

$$f_c(u)\,|_{max} = \frac{1}{0.863320\sqrt{2\pi}} \tag{5.9}$$

after normalization and other empirical modification on Equation (5.4), (5.5), and (5.6), the three fuzzy membership functions of **HPS**, **HPM**, and **HPC** soils are defined as

$$\mu_s(u) = \begin{cases} 1.0 & u>2.6575 \\ \exp\left(-\frac{1}{2}\left(\frac{u-2.6575}{0.834586}\right)^2\right) & u\le 2.6575 \end{cases} \tag{5.10}$$

$$\mu_m(u) = \exp\left(-\frac{1}{2}\left(\frac{u-1.35}{0.724307}\right)^2\right) \qquad -\infty < u < \infty \tag{5.11}$$

$$\mu_c(u) = \begin{cases} \exp\left(-\frac{1}{2}\left(\frac{u+0.1775}{0.86332}\right)^2\right) & u\ge 0.1775 \\ 1.0 & u<0.1775 \end{cases} \tag{5.12}$$

as shown in Figure 5.8.

The soil type **HPS**, **HPM**, and **HPC** with their fuzzy membership functions consist of the basic **CPT** fuzzy soil classification, which has a quite simple format. This classification has several characteristics and properties very useful. First, it is a real behavior type of soil classification in which not only the classification index is based

FUZZY MEMBERSHIP FUNCTIONS OF THREE SOIL TYPES

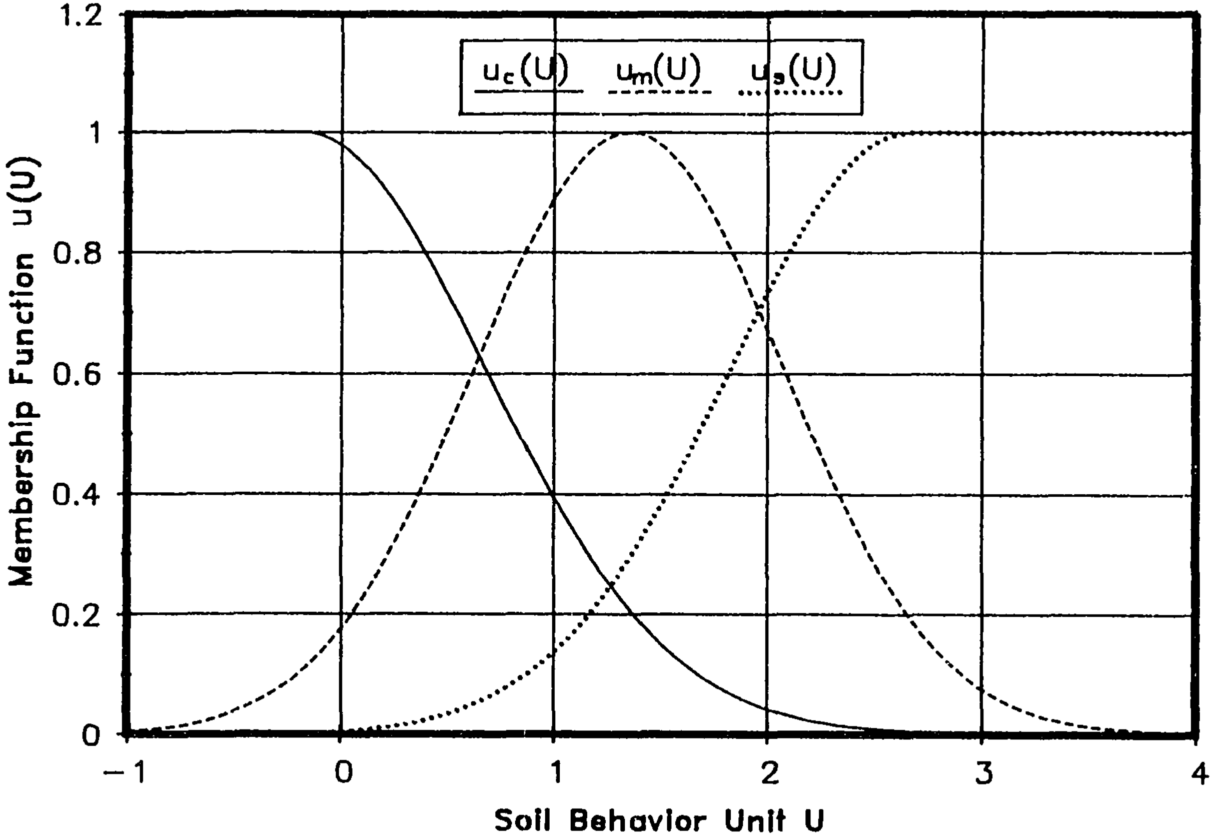


Figure 5.8 Tentative CPT Fuzzy Soil Classification Chart.

upon the measurements of soil behaviors but classification criteria are also really independent. Therefore, there is no uncertainty of randomness in this classification.

Second, soil types in this classification explicitly declare their relations with the corresponding compositional soil types. It is clearly defined that these soil types are the corresponding conventional ones of **USCS** with high probabilities but not always being of one hundred percent (100%). In this way, the uncertainty of randomness which exists in old **CPT** soil classifications has been put into the connotation of the concepts of the soil types. In other words, the emphasis has been put on the certainty of **CPT** soil classification. Since "high probability" is a fuzzy term, the changes of belonging to these soil types or the extensions of the concepts are described by the fuzzy membership functions.

Third, this classification has also included an empirical summarization of current knowledge about soil behaviors. The **HPS** type of soils generally has the properties of high strength, high permeability, and low compressibility, which will correspond to the higher tip resistance q_c and the lower friction ratio FR, therefore, the larger U values. The **HPC** type of soils then is supposed to have lower strength, lower permeability, and higher compressibility which are usually consistent with the lower q_c and higher FR, therefore, the lower U values. The engineering properties of **HPM** soils lie in between ones of the **HPS** and **HPC** soils. Thus, their U will take the values between the ones of **HPS** and **HPC** soils. All of these have been reflected clearly by this fuzzy soil classification.

Fourth, this classification separates the description of soil situation in-situ from the simplification and other decision makings on it. It tries to present the condition of soil types in-situ as real as it is. No conditional information truncations are performed right away. Later, if some criteria are required for certain actions, the conditional truncations on the fuzzy soil types can be executed accordingly. The resulting **C**risp **S**oil **T**ypes (**CST**) can be expressed by the fuzzy classification system with a group of α

$$CST_j = \{ u \mid \alpha_j \le \mu_i(u) \}, \quad i = c,\ m,\ s. \qquad (5.13)$$

Here i and j are the subscripts of the fuzzy and crisp soil types, separately. Therefore, as the supplement to the **CPT** fuzzy soil classification, a groups of level α values can also be given to users. These α values are determined by experts and will have different engineering concerns. In this way, users can have the choice of making their own decisions to serve their specific engineering purposes. Here is an example. Suppose that some **CPT** data are used to identify the liquefiable soil for some precautionary engineering measures. This soil can be recognized by the following criterion:

$$CST_{liquifiable} = \{ u \mid \alpha_{liquifiable} = 0.6 \le \mu_m(u) \}$$

Where $\mu_m(u)$ is given by Equation (5.11).

Last, this **CPT** fuzzy soil classification itself can serve the general communication purpose. An example of this can be: Given u = 2.0, the corresponding soil layer belongs to (might be more understandable to use the term of "looks like")

- a soil type of **HPC** with a degree of $\mu_c(2.0) = 0.04$,
- a soil type of **HPM** with a degree of $\mu_m(2.0) = 0.67$,
- a soil type of **HPS** with a degree of $\mu_s(2.0) = 0.73$,

depending upon which aspect you want to emphasize. This kind of description can be understood with the help of examples in our daily life. For instance, we can say that a boy looks like his mother with a degree o.6 (i.e. 60%) and his father with a degree of 0.7 (70%). People will accept such an introduction with little problem although it is a fuzzy description since human beings have the capacity to understand it. Also keep in mind that u = 2.0 is directly calculated from cone measurements of q_c and f_s. Therefore, the engineering behaviors of that soil layer is also determined.

5.7. CPT Soil Classification Package Suggested

A new package of **CPT** soil engineering classification based upon the cone tip resistance q_c and friction ratio FR is suggested now as a summary of the all previous analysis results. This new classification package should basically consist of the following procedures:

1). Transform a **CPT** sounding profile of parameters q_c and FR by Equation (2.1), (2.2), and (2.3) in order to get a corresponding profile of soil classification index, U;

2). Layer the index U profile by the ICC moving window method and then calculate the mean of U values for each layer to determine the soil behavior unit;

3). Predict the soil type of each layer by matching the soil behavior unit of that layer with some classification criteria.

Soil classification criteria, as one important result of this research, can be obtained from Table 4.2, 4.4, 4.5, and Figure 3.7, 3.8, 4.16, 4.17, 4.18, 4.19, 4.20, 4.21, 4.22, 4.23, and 5.8. These criteria are the indicators of an evolution process from the purely empirical to the purely theoretical. It will be the user's decision of which one to use. Also, if required, some other criteria from different considerations of soil classification can be merged with the ones suggested in this dissertation later. One important characteristic of this new **CPT** soil classification package is that the classification procedures suggested exactly follow the technical route taken in this research. Therefore, they have fully been explained and hopefully no ambiguous points will occur, which is what we initially want to obtain.

5.8. Summary

Soli classification has some inherent fuzziness due to its specific characteristics. These characteristics include **the comprehensive nature of the concepts of soil types** and **the law of quantity change to quality change** concerning soil composition and properties. They have to be treated properly in order to avoid possible confusion. Unfortunately, current classic classification methodology lacks this kind of capability. This situation, however, can be lessened by personal judgements of well trained and experienced engineers but even for them, the inconvenience caused by the fuzziness still exists. In an attempt to solve this problem, this chapter has explored the possibility and necessity to use a new classification approach which is based upon the fuzzy subset theory. It is concluded that this new technique has many advantages over traditional ones and it is an ideal tool to deal with the fuzziness. Therefore, it should be used when a new **CPT** soil classification is developed.

Apart from the inherent fuzziness, a **CPT** soil classification also has other characteristics. Specifically, it is basically a two dimensional problem and can be simplified into an one dimensional one if the approach suggested in Chapter 1 is used. On the other hand, the current practice has implicitly borrowed the classification criteria of **USCS** for **CPT** soil classifications, which has caused the overlaps in the resulting charts. Although such an uncertainty can be handled by a statistical approach as shown in Chapter 4, the results still can not make people feel totally comfortable. The problem is that the uncertainty of soil composition has been emphasized in these **CPT** soil classifications. This situation can be changed if the basic philosophy, not just the classification criteria, are borrowed from compositional soil classifications. Similar to a compositional soil classification, what are emphasized or relied on in a **CPT** soil classification should be certain.

A tentative **CPT** fuzzy soil classification is then suggested under all the above considerations. This new system basically consists of three types of soils with their fuzzy membership functions. These soil types are High Probability Clayey Soils (**HPC**), High Probability Mixed Soils (**HPM**), and High Probability Sandy Soils (**HPS**). The uncertainty of randomness discussed in Chapter 3 has been put into the connotation of the concepts of these new soil types. The resulting fuzziness is then described by the fuzzy membership functions, as shown in Equation (5.10), (5.11), (6.12), and Figure 5.8. These functions have included some up-to-date experience to determine fuzzy membership functions. They have empirically been decided by modifying the density functions of the corresponding compositional soil groups.

This new classification has several characteristics and properties very useful although its format is quite simple. First, it is a real soil classification of behavior type. Not only is its classification index based upon the measurements of soil behaviors but its classification criteria are also really independent. Then, this classification explicitly declares its relation with the compositional one. Its soil types correspond to the relevant compositional soil types with high probabilities but not one hundred percent (100%). Also, its soil types include an empirical summarization of current knowledge about soil behaviors. Furthermore, the classification has separated description from simplification and other decision makings. It tries to present the soil situation in-situ as real as it is and decision makings can later be implemented by conditional truncations. Finally, this classification can serve our general communication purposes.

Last in this chapter, a new package of **CPT** soil engineering classification based upon the cone tip resistance q_c and friction ratio FR is suggested and discussed as a summary of the all analysis results obtained in this research. An important characteristic of this new package is that the classification procedures exactly follows the technique route taken in this research. Therefore, they have fully been explained and hopefully no ambiguity will occur.

CHAPTER 6

DISCUSSION AND SUMMARY

The identification and classification of in-situ soil types are two aspects of a basic problem that an in-situ test usually needs to solve. This is true in the case of cone penetration technology. With the rapidly increasing usage of this technology in engineering practice, a demand for a new approach to classify soils based upon the results of in-situ cone (or piezocone) tests has occurred. This is because no soil samples are available in such kind of tests therefore the conventional soil classification procedure is not applicable. As a result, many soil classification charts based upon cone (or piezocone) testing data have been suggested and improved substantially.

However, there is an uncertainty in these soil classification charts. Soil types predicted by these charts sometimes do not match the real in-situ soil situation well. The following reasons have been found to be responsible for this phenomenon:

- The nature of soils is random;
- The correct usage of **USCS** is based upon expertise which varies individually. Therefore, if boring data from different sources are used to develop a **CPT** or **PCPT** soil classification, it will be possible to have some disparity on the classification results;
- Relations between compositional parameters and behavior measurements of soils are statistical or variational in nature;

- The test procedures in cone penetration technology are much more complicated than conventional laboratory soil classification index tests;
- Data from cone penetration are influenced by test equipments and procedures, soil compositions, and environments;
- The interpretation of data from cone penetration is mainly based on an empirical approach.

In order to perform a systematic research of this uncertainty problem, some preparative tasks have to be done in advance. These include finding an efficient soil classification index and a proper procedure for in-situ soil stratigraphy. These two issues are important since their results consist of the fundamental in a **CPT** or **PCPT** soil classification. An efficient soil classification index instead of a current two dimensional chart format of a **CPT** soil classification will well preserve the advantage of continuously describing in-situ soil conditions. Such an index can also greatly reduce the difficulty and complexity of a possible theoretic analysis. Similarly, a proper procedure for in-situ soil stratigraphy will pave the way to find out the representative correlations between soil composition and soil behavior due to cone or piezocone penetration.

It has been observed that there are two tendencies in most existing **CPT** or **PCPT** soil classification charts, where soil type changes in one direction and in-situ soil state changes in another. As the results of the research work presented in Chapter 1, two-dimensional soil classification charts can be simplified into one-dimensional soil classification systems by a linear conformal mapping. This conformal mapping is composed of Equation (1.3) and (1.4). Accordingly, Douglas et al's chart (1981), Olsen et al's chart (1988), and P. K. Robertson's charts (1990), as shown in Figure 1.1, 1.2, and 1.3, have been transformed into their corresponding U --- V planes as shown in Figure 1.6, 1.8, 1.10, 1.12. The specific formulas for them are Equation (1.6), (1.7), (1.8), (1.9), (1.10), (1.11), (1.12), (1.13), (1.14), and (1.15). Based upon these transformations, the corresponding one-dimensional soil classification systems can be established through the soil classification index, U, and the new systems are at least equivalent to the original ones.

It is clear that the generality of this simplifiability on current cone or piezocone soil classification charts reflects an important fact. Cone tip resistance q_c, friction ratio FR, and pore water pressure u are not fundamental parameters for soil classification even if they are normalized with respect to some factors. Therefore, reorganizing these parameters to obtain a primary indicator of soil types should be an essential step to develop an efficient **CPT** soil classification system. This can be implemented by the conform mapping exhibited in Chapter 1. One benefit of this reorganization is that the advantage of continuous description of site condition in raw cone data is well preserved in the analysis results of site stratigraphy, as shown in Figure 1.13 and 1.14. Consequently, the efficiency of expressing the interpretation of cone data is greatly improved. Also the dimension of resulting new classifications is reduced to one. This left parameter is the soil classification index, U, determined by Equation (1.3). Therefore, subsequent analyses on **CPT** testing data concerning soil classification will mainly be based upon the profiles of soil classification index, U.

Concerning the in-situ soil stratigraphy, it is known that in the traditional site investigation, the identification of soil layers should be done first and the classification of soil types are next due to the characteristics of boring technique. Unfortunately, this pattern has little change in the interpretation of cone testing data for site investigation due to the characteristics of cone penetration. These characteristics have made the concept of soil behavior unit imperative in a **CPT** soil classification process. Such a concept has in fact reflected the average responses of soils to cone penetration over a layer and can be implemented with the help of a moving window approach. In this new approach, a statistic profile of Intraclass Correlation Coefficient (**ICC**), ρI, can be obtained, as shown in Figure 2.1. This coefficient is defined by Equation (2.8) and calculated by Equation (2.10). As a result, a set of soil layers over a U profile can be determined accordingly and the corresponding average U values are obtained. Each soil layer therefore will have a property of statistical homogeneity. In such a way, soil behavior units for those soil profiles can be decided objectively and their corresponding soil types can be determined by the related boring data given. This process of data manipulation is called the preliminary data reduction and its product is a set of soil behavior units with corresponding soil type information, which will pave the way for establishing the correlation between soil behaviors and soil types.

The preliminary data reduction is accomplished over the **CPT** in-situ data from eight (8) testing sites, as shown in Table 3.1. Some empirical correlations between soil types and soil behavior units are obtained accordingly, as demonstrated in Figure 3.5, 3.7, and 3.8. The results also confirm the uncertainty discussed at beginning of this dissertation. They indicate that in general, different soil types have different in-situ behaviors but they sometimes can have similar ones. Current **CPT** soil classifications try to make use of the first fact but the second one has caused the overlaps in their charts. If the simplification procedure and the concept of soil behavior unit suggested in this dissertation are adopted, these overlaps will in general be independent on the depths and thicknesses of soil layers, as shown in Figure 3.1 and 3.2. Consequently, no corrections based upon them can reduce the uncertainty in these charts.

It is obviously not enough to only know the existence of the uncertainty. The knowledge of the probabilities with which the overlaps will occur is more important. Therefore, some statistical models are suggested in order to find out some ways to determine these probabilities. Actually, two methods of predicting soil types: Region Estimation and Point Estimation, as shown in Table 4.2, 4.4, 4.5, Figure 4.16, 4.17, 4.18, 4.19, 4.20, 4.21, 4.22, and 4.23, are developed to quantitatively describe these overlaps among different soil types in a **CPT** soil classification. All the results are based upon the normal distribution assumption of soil behavior unit, U, of seven (7) types of soils. The validation of the assumption is checked by Chi-square test and Kolmogorov-Smirnov one-sample test. The results of those tests indicate that this assumption is valid with a high confidence for most soil types except for soil type CH and SM. It is not clear right now why CH and SM samples do not follow a normal distribution well. More CPT data are needed to examine and explain this phenomenon and to verify and validate the distributions of soil classification index, U, for other soil types.

Region Estimation is generally similar to conventional soil classifications to classify soils. The predicted soil type depends upon the region a soil sample falls in. The difference is that in the new method, each region will correspond not to just one type of soil but several ones with different possibilities. It appears that the new method will make the soil classification complicated, but it is real life. The basic philosophy of the method is to try to keep and present the original information on soil types as it is and let users make their decision of simplification.

The problem for Region Estimation is that different points in each region are supposed to have exactly the same statistical property so that they are treated in exactly the same way. Unfortunately, this assumption is not true in most cases since different points in a region will have different probabilities to belong to (receive) different types of soils. This problem can be solved by Point Estimation which classifies soils directly by probability. In this method, every point along the U axis is treated differently. The basic question this new approach intends to answer is: given a specific value of soil in-situ behavior unit, U, what are the probabilities with which the corresponding soil sample belongs to different types of soils? The whole method is based upon a probabilistic model of two dimensions, one of which is the continuous random variable of soil classification index, U, representing soil in-situ behavior units. The other is a discrete random variable representing soil types. Therefore, if a **CPT** soil behavior unit, U, is given, both Region and Point Estimation can be performed accordingly.

Apart from the uncertainty of randomness discussed previously due to the lack of the law of causation, soil classification also has another kind of uncertainty called fuzziness. This uncertainty is associated with the lack of the law of exclude middle. It is inherent in soil classification because of some specific characteristics of the problem. These characteristics include **the comprehensive nature of the concepts of soil types** and **the law of quantity change to quality change** concerning soil composition and properties. This fuzziness has to be treated properly in order to avoid possible confusion. Unfortunately, current classic classification methodology lacks the capability to handle this kind of uncertainty. This situation, however, can be lessened by personal judgements of well trained and experienced engineers, but even for them, the inconvenience caused by the fuzziness still exists. In an attempt to solve this problem, Chapter 5 has explored the possibility and necessity to use the fuzzy subset theory. It is concluded that the fuzzy subset theory is an ideal tool to deal with the uncertainty of fuzziness in soil classification problem. The resulting fuzzy soil classification will have many advantages over traditional ones. Therefore, the new technique should be used when a **CPT** soil classification is developed.

A **CPT** soil classification also has its own characteristics. Specifically, it is basically a two dimensional problem and can be simplified into an one dimensional one if the approach suggested in Chapter 1 is used. On the other hand, current practice has implicitly borrowed the soil classification criteria of **USCS** for **CPT** soil classifications, which has caused the overlaps in the resulting charts. Although such an uncertainty can be handled by a statistical approach as shown in Chapter 4, the results still can not make people feel totally comfortable. The problem is that the technique approach adopted so far has actually put an emphasis on the uncertainty: soil composition.

The situation can be changed if the basic philosophy, not just the classification criteria, are borrowed from the compositional soil classifications. Just like what was done in a compositonal soil classification, the certainty of soil behavior and other useful information should be emphasized in a **CPT** soil classification. A tentative **CPT** fuzzy soil classification is then suggested under all the above considerations. This new system basically consists of three types of soils with their fuzzy membership functions. These soil types are High Probability Clayey Soils (**HPC**), High Probability Mixed Soils (**HPM**), and High Probability Sandy Soils (**HPS**). The uncertainty of randomness discussed in Chapter 3 has been put into the connotation of the concepts of these new soil types. The resulting fuzziness is then described by the fuzzy membership functions, as shown in Equation (5.10), (5.11), (6.12), and Figure 5.8. These functions have included some up-to-date experience to determine fuzzy membership functions and are empirically decided by modifying the density functions of the corresponding compositional soil groups.

This new classification has several characteristics and properties very useful although its format is quite simple. First, it is a real soil classification of behavior type. Not only is its classification index based upon the measurements of soil behaviors but its classification criteria are also really independent. Then, this classification explicitly declares its relation with the compositional one. Its soil types correspond to the relevant compositional soil types with high probabilities but not one hundred percent (100%). Also, its soil types include an empirical summarization of current knowledge about soil behaviors. Furthermore, the classification has separated description from simplification and other decision makings. It tries to present the soil situation in-situ as real as it is and decision makings can later be implemented by conditional truncations. Finally, this classification can serve our general communication purposes.

Finally at the end of Chapter 5, a new package of **CPT** soil engineering classification based upon the cone tip resistance q_c and friction ratio FR is suggested and discussed as a summary of the all analysis results obtained in this research. This new classification package consists of following procedures:

1). Transform a **CPT** sounding profile of parameter q_c and FR by using Equation (2.1), (2.2), and (2.3) to get a corresponding profile of soil classification index, U;

2). Layer the index U profile by the ICC moving window method and then calculate the mean of U values for each layer to determine the soil behavior unit;

3). Predict the soil type of each layer by matching the soil behavior unit, U, of that layer with some classification criteria.

Soil classification criteria, as one important result of this research, can be obtained from Table 4.2, 4.4, 4.5, and Figure 3.7, 3.8, 4.16, 4.17, 4.18, 4.19, 4.20, 4.21, 4.22, 4.23, and 5.8. These criteria are the indicators of an evolution process from the purely empirical to the purely theoretical. Users will have the choice of which one to use. Also, if required, some other criteria from different considerations of soil classification can be merged with the ones suggested in this dissertation later. An important characteristic of this new package is that the classification procedures exactly follows the technique route taken in this research. Therefore, they have fully been explained and hopefully no ambiguity will occur.

Concerning the further research effort on the soil engineering classification of cone penetration technology, it is clear that more reliable cone testing data are needed to verify and validate the density functions of the soil classification index U (soil behavior unit) for each type of soil obtained in this research. If this is accomplished, the suggested procedures to predict soil types in this dissertation will have a more sound foundation to be used in engineering practice. Also, the basic philosophy and technical approaches adopted in this research can be used to develop a **PCPT** soil classification package although some modification might be necessary. Most importantly, in order to make mature the fuzzy **CPT** soil classification suggested in this dissertation, more experts' expertise should be collected and used.

CHAPTER 7

CONCLUSIONS

A systematic investigation on **CPT** soil engineering classification has been accomplished and following conclusions are available:

1). There are two types of uncertainties, randomness and fuzziness, in a **CPT** soil engineering classification.

2). Several reasons can be found for the uncertainty of randomness in a **CPT** soil engineering classifications. They are:

- The nature of soils is random;
- The correct usage of **USCS** is based upon expertise which varies individually. Therefore, if boring data from different sources are used to develop a **CPT** soil engineering classification, it will be possible to have some disparity on the classification results;
- Relations between compositional parameters and behavior measurements of soils are statistical or variational in nature;
- The test procedures in cone penetration technology are much more complicated than conventional laboratory soil classification index tests;
- Data from cone penetration are influenced by test equipments and procedures, soil compositions, and environments;

- The interpretation of data from cone penetration is mainly based on an empirical approach.

3). The fuzziness in a **CPT** soil engineering classification is mainly due to the comprehensive nature of the concepts of soil types and the quantity change of soil composition and properties to quality change of soils.

4). The suggested conformal mapping is useful and the soil classification index, U, obtained from this mapping can be adopted as an indicator of soil types in a statistical sense.

5). The "in-situ soil behavior unit" is a necessary concept based upon which layering soil profiles and predicting soil types can be properly integrated together.

6). The suggested **ICC** moving window method is a proper technique to objectively obtain statistically homogeneous layers.

7). One of the necessary conditions to eliminate the obstacle of the uncertainty of randomness from a **CPT** soil engineering classification is to adopt independent soil classification criteria.

8). The fuzzy **CPT** soil engineering classification suggested is an ideal system to solve all the problems met in a **CPT** soil engineering classification.

9). A new package of **CPT** soil engineering classification is suggested and this package consists of three procedures, i.e., transforming, layering, and classifying. All the procedures together organically comprise the whole new soil classification.

10). Several sets of **CPT** soil engineering classification criteria are recommended from this investigation. They are the indicators of an evolution process from the purely empirical to the purely theoretical.

11). More **CPT** and related boring data with good quality are needed to improve the reliability of the new package of **CPT** soil engineering classification suggested.

REFERENCES

ASTM Designation: D 2487 - 85, "Standard Test Method for Classification of Soils for Engineering Purposes", Annual Book of ASTM Standards, Soil and Rock; Building Stones, Section 4, volume 04.08, 1987.

ASTM Designation: D 2488 - 84, "Standard Practice for Description and Identification of Soils (Visual-Manual Procedure)", Annual Book of ASTM Standards, Soil and Rock; Building Stones, Section 4, volume 04.08, 1987.

ASTM Designation: D 3282 - 83, "Standard Practice for Classification of Soils and Soil-Aggregate Mixtures for Highway Construction Purposes", Annual Book of ASTM Standards, Soil and Rock; Building Stones, Section 4, volume 04.08, 1987.

Begemann, H.K., "The Friction Jacket Cone as an Aid in Determining the Soil Profile", Proceedings 6th International Conference on Soil Mechanics and Foundation Engineering, Vol. 1, pp 17-20, 1965.

Blockley, D.I., "Written Comments", Proceedings of NSF Workshop on Civil Engineering Applications of Fuzzy sets, Purdue University, West Lafayette, pp 180-181, 1985.

Brown, C.B., Chameau, J.L., Palmer, R., and Yao, J.T.P., Proceedings of NSF Workshop on Civil Engineering Applications of Fuzzy sets, Purdue University, West Lafayette, 1985.

Campanella R. G. and Robertson P. K. 1988., "Current Status of the Piezocone Test". Proc. Penetration Testing 1988, ISOPT-1, De Ruiter(ed.), Orlando. Vol.I: pp 93-116, 1988.

Casagrande, A., "Classification and Identification of Soils," Transactions, ASCE, Vol. 113, pp 901-991, 1948.

Chameau, J.L. and Santamarina, J.C., "Knowledge-Based System for Soil Improvement", ASCE, ISSN 0887-3801/89/0003-0253, Paper No. 23647.

Cheng-hou, Z., Greeuw, G., Jekel, J., and Rosenbrand, W., "A New Classification Chart for Soft Soils Using the Piezocone Test", Engineering Geology, 29, pp 31-47, 1990.

De Beer, E. E., Goelen, E., Heynen, W. J., and Joustra, K., ISSMFE Technical Committee on Penetration Testing CPT Working Party, "Cone Penetration Test (CPT): International Reference Test Procedure", Proc. Penetration Testing 1988, ISOPT-1, De Ruiter(ed.), Orlando. Vol.I: pp 27-51, 1988.

Douglas, B. J. and Olsen, R. S., "Soil Classification Using the Electric Cone Penetrometer". Proc. Cone Penetration Testing and Experience, ASCE, St. Louis, MI:209-227, 1981.

Gunaratne, M., Chameau, J.L., and Altschaeffl, A.G., "Fuzzy Multi-Attribute Decision Making in Pavement Management", Civil Engineering System, Vol. 2, September, 1985.

Hoel, P.G., Port, S.C., and Stone, C.J., Introduction to Statistical Theory, Houghton Mifflin Company, Boston, 1971.

Holtz, R. D. and Kovacs, W. D., An Introduction to Geotechnical Engineering. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1981.

Jones, G. A. and Rust, E. A., "Piezometer Penetration Testing CUPT", Proceedings of the 2nd European Symposium on Penetration Testing , ESOPT II, Amsterdam, Vol.2, pp 607-613, 1982.

Kacprzyk, J., Multistage Decision-Making under Fuzziness, Verlag TUV Rheinland, 1983.

Kandel, A., Fuzzy Mathematical techniques with Applications, Addison-Wesley Publishing Company, Inc., 1986.

Kaufmann, A., Introduction to the Theory of Fuzzy Subsets, Volume I, Fundamental Theoretical Elements, Academic Press, Inc., New York, 1975.

Kaufmann, A. and Gupta, M.M., Fuzzy Mathematical Models in Engineering and Management Science, Elsevier Science Publishers Company Inc., New York, 1988.

Kaufmann, A. and Gupta, M.M., Introduction to Fuzzy Arithmetic, Theory and Applications, Van Nostrand Reinhold Company Inc., New York, 1985.

Kelly, P. J. and Straus E. G., Elements of Analytic Geometry and Linear Transformations. Scott, Foresman and Company, 1968.

Kreyszig, E., Advanced Engineering Mathematics, Sixth edition, John Wiley & Sons, Inc., New York, 1988.

Kruse, R. and Meyer, K.D., Statistics with Vague Data, D. Reidil Publishing Company, Dordrecht, Holland, 1987.

Larson, N. B. and Mitchell, B., "Cone Penetrometer Use on Uranium Mill Tailings", USE OF IN SITU TESTS IN GEOTECHNICAL ENGINEERING, Proc. INSITU 86, Geotechnical Special Publication, No. 6, ASCE, Blacksburg, VA: pp 700-713, 1986.

Lima, D. C. and Tumay, M. T., "Scale Effects in Cone Penetration Tests". ASCE Geotechnical Engineering Congress 1991, No.27, Vol I: pp 38-51, 1991.

Meigh, A.C., Cone Penetration Testing Methods and Interpretation, CIRIA Ground Engineering Report: In-situ Test, Butterworths, 1987.

Norris, G.M. and Holtz, R.D., Cone Penetration Testing and Experience, ASCE, New York, 1981.

Olsen, R. S. and Farr, J. V., "Site characterization Using the Cone Penetrometer Test", USE OF IN SITU TESTS IN GEOTECHNICAL ENGINEERING, Proc. INSITU 86, Geotechnical Special Publication, No. 6, ASCE, Blacksburg, VA: pp 864-868, 1986.

Olsen, R. S. and Malone, P. G., "Soil Classification and Site Characterization Using the Cone Penetrometer Test". Proc. Penetration Testing 1988, ISOPT-1, De Ruiter(ed.), Orlando. Vol.II: pp 887-893, 1988.

Robertson, P.K. and Campanella, R.G., "Guidelines for Use & Interpretation of the Electronic Cone Penetration Test", Hogentohler & Company, Inc., 1984.

Robertson, P. K., Campanella, R. G., Gillespie, D., and Grieg, J., "Use of Piezometer Cone Data", USE OF IN SITU TESTS IN GEOTECHNICAL ENGINEERING, Proceedings of In-Situ '86, Geotechnical Special Publication, No. 6, ASCE, Blacksburg, Virginia, pp 1263--1280, 1986.

Robertson, P.K., "Soil Classification Using the Cone Penetration Test," Canadian Geotechnical Journal, 27, pp 151-158, 1990.

Ruiter, J.D., Penetration Testing 1988, ISOPT-1, Volume 1, 2, A.A. Balkema, 1988.

Sanglerat, G., The Penetrometer and Soil Exploration, Elsevier Scientific Publishing Company, New York, 1972.

Santamarina, J.C., Fuzzy Sets and Knowledge Systems in Geotechnical Engineering, a dissertation for PhD, Purdue University, 1987.

Santamarina, J.C. and Chameau, J.L., "Fuzzy Windows and Classification System", Int. J. Man-Machine Studies (1990) 32, pp 187-201, 1990.

Santamarina, J.C. and Chameau, J.L., "Expert Systems for Geotechnical Engineers", ASCE, ISSN 0887-3801/87/004-0241, Paper No. 21897, 1987.

Schmertmann, J.H., Guidelines for Cone Penetration Test, Performance and Design, U.S. Department of Transportation, Federal Highway Administration Report NO. FHWA-TS-78-209, 1978.

Senneset, K. and Janbu, N., "Shear Strength Parameters Obtained from Static Cone Penetration Tests", ASTM STP 883, Symposium, San Diego, 1984.

Senneset, K., Sandven, R., and Janbu, N., "The Evaluation of Soil Parameters from Piezocone Tests", Paper No. 88-0467, Transportation Research Board, 68th Annual Meeting, Washington, D.C., January pp 22-26, 1989.

Smith, G.N., Probability and Statistics in Civil Engineering --- an Introduction, Nichols Publishing Company, New York, 1986.

Souflis, C. and Grivas, D.A., "Fuzzy Set Approach to Linguistic Seismic Load and Damage Assessments", ASCE, ISSN 0733-9399/86/0006-0605, Paper No. 20699, 1986.

STATGRAPHICS, Version 4.0, STATISTICAL GRAPHICS SYSTEM BY STATISTICAL GRAPHICS CORPORATION, STSC, Inc., 1989.

Tumay, M.T., Field Calibration of Electric Cone Penetrometers in Soft Soil - Executive Summary, Report No.FHWA/LA/LSU-GE-85/2, U.S. Department of Transportation, Federal Highway Administn., 1985.

USGS Open-File Report No. 81-284, Evaluation of the Cone Penetrometer for Liquefaction Hazard Assessment, Prepared by Fugro, Inc., 1980.

Waterways Experiment Station, The Unified Soil Classification System, Technical Memorandum NO. 3-357, Vicksburg, Mississippi, 1951.

Webster, R., Quantitative and Numerical Methods in Soil Classification and Survey, Clarendon Press, Oxford, 1977.

Wickremesinghe, D. S., Statistical Characterization of Soil Profiles Using In-Situ Tests, a Ph. D. Dissertation, Department of Civil Engineering, University of British Columbia, 1989.

Zadeh, L. A., "Fuzzy Set", Information and Control 8, pp 338-353, 1965.

APPENDIX A

BASIC FORMULAS FOR A CONFORMAL TRANSFORMATION

Given:

$z_1 = x_1 + y_1 i, \qquad z_2 = x_2 + y_2 i, \qquad z_3 = x_3 - y_3 i;$

$w_1 = u_1 + v_1 i, \qquad w_2 = u_2 + v_2 i, \qquad w_3 = u_3 + v_3 i.$

$$w = f(z) = u(x,y) + iv(x,y), \qquad z = x + iy \tag{A.1}$$

$$\frac{w - w_1}{w - w_3} \, \frac{w_2 - w_3}{w_2 - w_1} = \frac{z - z_1}{z - z_3} \, \frac{z_2 - z_3}{z_2 - z_1} \tag{A.2}$$

Find u(x,y) and v(x,y) in Equation (A.1).

In Equation (A.2), assume

$$S = \frac{z_2 - z_3}{z_2 - z_1} \, \frac{w_2 - w_1}{w_2 - w_3} = s_1 + s_2 i \tag{A.3}$$

We get

$$\frac{w - w_1}{w - w_3} = \frac{z - z_1}{z - z_3} \cdot S \tag{A.4}$$

Therefore

$$w - w_1 = \frac{z - z_1}{z - z_3} \, S \, (w - w_3) \tag{A.5}$$

or

$$w\left(1 - \frac{z - z_1}{z - z_3} \cdot S\right) = w_1 - \frac{z - z_1}{z - z_3} \cdot S \cdot w_3 \tag{A.6}$$

After some simplification, the w will be

$$w = \frac{z(w_1 - Sw_3) + z_1 S w_3 - w z_3}{z(1 - S) + z_1 S - z_3} \tag{A.7}$$

Assume

$$A = w_1 - Sw_3 = a_1 + a_2 i \tag{A.8}$$

$$B = z_1 S w_3 - w_1 z_3 = b_1 + b_2 i \tag{A.9}$$

$$C = 1 - S = c_1 + c_2 i \tag{A.10}$$

$$D = z_1 S - z_3 = d_1 + d_2 i \tag{A.11}$$

Plug them back into Equation (A.7), the w can be written as

$$w = \frac{(x + yi)\ (a_1 + a_2 i) + b_1 + b_2 i}{(x + yi)\ (c_1 + c_2 i) + d_1 + d_2 i} \tag{A.12}$$

or

$$w = \frac{a_1 x - a_2 y + b_1 + i\,(a_2 x + a_1 y + b_2)}{c_1 x - c_2 y + d_1 + i\,(c_2 x + c_1 y + d_2)} \tag{A.13}$$

By some derivation, we get

$$w = \frac{(a_1x - a_2y + b_1)\ (c_1x - c_2y + d_1)}{(c_1x-c_2y+d_1)^2 + (c_2x + c_1Y + d_2)} +$$

$$+ \frac{(a_2x + a_1y + b_2)\ (c_2x + c_1y + d_2)}{(c_1x-c_2y+d_1)^2 + (c_2x + c_1Y + d_2)} +$$

$$+ i\frac{(c_1x - c_2y + d_1)\ (a_2x + a_1y + b_2)}{(c_1x-c_2y+d_1)^2 + (c_2x + c_1Y + d_2)} -$$

$$- i\frac{(c_2x + c_1y + d_2)\ (a_1x - a_2y + b_1)}{(c_1x-c_2y+d_1)^2 + (c_2x + c_1Y + d_2)} \qquad \text{(A.14)}$$

That is

$$u = \frac{(a_1x - a_2y + b_1)\ (c_1x - c_2y + d_1)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} +$$

$$+ \frac{(a_2x + a_1y + b_2)\ (c_2x + c_1y + d_2)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} \qquad \text{(A.15)}$$

$$v = \frac{(c_1x - c_2y + d_1)\ (a_2x + a_1y + b_2)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} -$$

$$- \frac{(a_1x - a_2y + b_1)\ (c_2x + c_1y + d_2)}{(c_1x - c_2y + d_1)^2 + (c_2x + c_1y + d_2)^2} \qquad \text{(A.16)}$$

These two formulas are the basic ones for the conformal transformation we are trying to find out. Their coefficients can be determined as follows.

According to Equation (A.3) and (A.8), the complex number $A = a_1 + a_2 i$ will be

$$a_1 + a_2 i = u_1 + v_1 i - (s_1 + s_2 i)\ (u_3 + v_3 i)$$

$$= u_1 - s_1 u_3 + s_2 v_3 + i\,(v_1 - s_1 v_3 - s_2 u_3) \qquad \text{(A.17)}$$

Therefore, a_1 and a_2 will be

$$a_1 = u_1 - s_1 u_3 + s_2 v_3 \qquad \text{(A.18)}$$

$$a_2 = v_1 - s_1 v_3 - s_2 u_3 \qquad \text{(A.19)}$$

Similarly, from Equation (A.9),

$$B = -(u_1 + v_1 i)\ (x_3 + y_3 i) + (u_3 + v_3 i)\ (x_1 + y_1 i)\ (s_1 + s_2 i)$$

$$= v_1 y_3 - u_1 x_3 + u_3 (x_1 s_1 - y_1 s_2) - v_3 (x_1 s_2 + y_1 s_1) +$$

$$+ i\ (u_3 (x_1 s_2 + y_1 s_1) + v_3 (x_1 s_1 - y_1 s_2) - (v_1 x_3 + u_1 y_3)) \qquad \text{(A.20)}$$

Thus, b_1 and b_2 are

$$b_1 = v_1 y_3 - u_1 x_3 + u_3 (x_1 s_1 - y_1 s_2) - v_3 (x_1 s_2 + y_1 s_1) \qquad \text{(A.21)}$$

$$b_2 = u_3 (x_1 s_2 + y_1 s_1) + v_3 (x_1 s_1 - y_1 s_1) - (v_1 x_3 + u_1 y_3) \qquad \text{(A.22)}$$

In a same way, from Equation (A.10), c_1 and c_2 are

$$c_1 = 1 - s_1 \tag{A.23}$$

$$c_2 = - s_2 \tag{A.24}$$

Since, from Equation (A.11),

$$D = -x_3 - y_3 i + (x_1 + y_1 i)\ (s_1 + s_2 i)$$

$$= x_1 s_1 - y_1 s_2 - x_3 + i\,(s_1 y_1 + x_1 s_2 - y_3) \tag{A.25}$$

Consequently,

$$d_1 = x_1 s_1 - y_1 s_2 - x_3 \tag{A.26}$$

$$d_2 = s_1 y_1 + x_1 s_2 - y_3 \tag{A.27}$$

Equation (A.18), (A.19), (A.21), (A.22), (A.23), (A.24), (A.26), and (A.27) are the formulas to calculate the coefficients in Equation (A.15) and (A.16). Suppose the following values of z and w are taken,

$z_1 = 0.63 + i$, $\quad z_2 = 1.34$, $\quad z_3 = 3 - 0.73\ i$

$w_1 = -5$, $\quad w_2 = 0$, $\quad w_3 = 5$

According to Equation (A.3),

$$S = \frac{1.34 - 3 + 0.73i}{1.34 - 0.63 - -}\ \frac{5}{-5} = \frac{1.66 - 0.73i}{0.71 - i}$$

$$= \frac{(1.66 - 0.73i)(0.71 + i)}{0.71^2 + 1^2} = \frac{1.9086}{1.5041} + \frac{1.1417}{1.5041} i$$

Therefore, from Equation (A.18) and (A.19),

$$a_1 = -5 - 5\frac{1.9086}{1.5041} = -11.3446$$

$$a_2 = -0.7590 \cdot 5 = -3.7953$$

and from Equation (A.21) and (A.22),

$$b_1 = 5 \cdot 3 + 5(0.63 \cdot 1.2689 - 1 \cdot 0.7590) = 15.2019$$

$$b_2 = 5(0.63 \cdot 0.7590 + 1 \cdot 1.2689) + 5(-0.73) = 5.085$$

Also, from Equation (A.23) and (A.24),

$$c_1 = 1 - 1.2689 = -0.2689$$

$$c_2 = - s_2 = - 0.7590$$

Finally, based upon Equation (A.26) and (A.27),

$$d_1 = 0.63 \cdot 1.2689 - 1 \cdot 0.7590 - 3 = -2.9596$$

$$d_2 = 1.2689 \cdot 1 + 0.63 \cdot 0.7590 + 0.73 = 2.4771$$

These are the coefficients of Equation (A.15) and (A.16) under the given conditions.

APPENDIX B

ANOTHER FORMULA FOR THE BETWEEN CLASS VARIANCE Y_b^2

According to the definitions, we have

$$\gamma_b^2 = \frac{1}{2n-1} \sum_{i=1}^{n} [(y_i - \mu)^2 + (z_i - \mu)^2] \tag{B.1}$$

$$\mu = \frac{1}{2n} \sum_{i=1}^{n} (y_i + z_i) \tag{B.2}$$

here, the y_i and z_i are the sample readings on each side of the middle line in a moving window, separately. The μ can also be written as,

$$\mu = \frac{n}{2n} \left(\frac{1}{n} \sum_{i=1}^{n} y_i + \frac{1}{n} \sum_{i=1}^{n} z_i\right) = \frac{1}{2} (\mu_1 + \mu_2) \tag{B.3}$$

Since

$$\sum_{i=1}^{n} (y_i - \mu)^2 + \sum_{i=1}^{n} (z_i - \mu)^2 =$$

$$= \sum_{i=1}^{n} (y_i^2 - 2\mu y_i + \mu^2 + z_i^2 - 2\mu z_i + \mu^2)$$

$$= \sum_{i=1}^{n} (y_i^2 + z_i^2) - 2\mu \sum_{i=1}^{n} (y_i + z_i) + 2n\mu^2$$

$$= \sum_{i=1}^{n} (y_i^2 + z_i^2) - (\mu_1 + \mu_2) n \left(\frac{1}{n}\sum_{i=1}^{n} y_i + \frac{1}{n}\sum_{i=1}^{n} z_i\right) + \frac{n}{2}(\mu_1 + \mu_2)^2$$

$$= \sum_{i=1}^{n} (y_i^2 + z_i^2) - n(\mu_1 + \mu_2)^2 + \frac{n}{2}(\mu_1 + \mu_2)^2$$

$$= \sum_{i=1}^{n} (y_i^2 + z_i^2) - \frac{n}{2}(\mu_1 + \mu_2)^2$$

$$= \sum_{i=1}^{n} y_i^2 - 2n\mu_1^2 + n\mu_1^2 + \sum_{i=1}^{n} z_i^2 - 2n\mu_2^2 +$$

$$+ n\mu_2^2 + n\mu_1^2 + n\mu_2^2 - \frac{n}{2}(\mu_1 + \mu_2)^2$$

$$= \sum_{i=1}^{n} y_i^2 - 2\mu_1 \sum_{i=1}^{n} y_i + \sum_{i=1}^{n} \mu_1^2 + \sum_{i=1}^{n} z_i^2 - 2\mu_2 \sum_{i=1}^{n} z_i + \sum_{i=1}^{n} \mu_2^2 +$$

$$+ n\mu_1^2 + n\mu_2^2 - \frac{n}{2}(\mu_1 + \mu_2)^2$$

$$= \sum_{i=1}^{n} (y_i - \mu_1)^2 + \sum_{i=1}^{n} (z_i - \mu_2)^2 + \frac{n}{2}\mu_1^2 + \frac{n}{2}\mu_2^2 - n\mu_1\mu_2$$

$$= \left[\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \mu_1)^2 + \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \mu_2)^2\right](n-1) + \frac{n}{2}(\mu_1 - \mu_2)^2$$

$$= (n - 1)\ (\sigma_1^2 + \sigma_2^2) + \frac{n}{2}(\mu_1 - \mu_2)^2$$

Therefore, we have

$$Y_b^2 = \frac{n - 1}{2n - 1}(\sigma_1^2 + \sigma_2^2) + \frac{n}{2(2n - 1)}(\mu_1 - \mu_2)^2 \qquad \text{(B.4)}$$

APPENDIX C

FORMULA TO CALCULATE THE BOUNDARY VALUES AMONG DIFFERENT SOIL CLASSIFICATION REGIONS

Given

$$1 - F_i(U_{i,i+1}) = F_{i+1}(U_{i,i+1}) \tag{C.1}$$

Find

$$U_{i,i+1} = \frac{\mu_{i+1}\sigma_i + \mu_i\sigma_{i+1}}{\sigma_i + \sigma_{i+1}} \tag{C.2}$$

here, μ_i and σ_i are the mean and standard deviation of the normal density function of F_i and μ_{i+1} and σ_{i+1} are the mean and standard deviation of the normal density function of F_{i+1}, separately.

According to the definition of F(U), Equation (C.1) can be written as

$$1 - \int_{-\infty}^{U_{i,i+1}} f_i(u)\,du = \int_{-\infty}^{U_{i,i+1}} f_{i+1}(u)\,du \tag{C.3}$$

or

$$\int_{U_{i,i+1}}^{\infty} f_i(u)\,du = \int_{-\infty}^{U_{i,i+1}} f_{i+1}(u)\,du \tag{C.4}$$

Here, $f_i(u)$ and $f_{i+1}(u)$ are the density functions assumed with a normal form. Their means and standard deviations are μ_i, σ_i and μ_{i+1}, σ_{i+1}, separately. Therefore, Equation (C.4) can be rewritten as

$$-\int_{\infty}^{\frac{U_{i,i+1}-\mu_i}{\sigma_i}} \Phi(u)\,du = \int_{-\infty}^{\frac{U_{i,i+1}-\mu_{i+1}}{\sigma_{i+1}}} \Phi(u)\,du \tag{C.5}$$

in which $\Phi(U)$ is the density of standardized normal distribution.

Since

$$-\int_{\infty}^{0} \Phi(u)\,du = \int_{-\infty}^{0} \Phi(u)\,du = \frac{1}{2} \tag{C.6}$$

Equation (C.5) can be simplified as

$$-\int_{0}^{\frac{U_{i,i+1}-\mu_i}{\sigma_i}} \Phi(u)\,du = \int_{0}^{\frac{U_{i,i+1}-\mu_{i+1}}{\sigma_{i+1}}} \Phi(u)\,du \tag{C.7}$$

Also, due to $\Phi(u) = \Phi(-u)$, the left side (L.S.) of Equation (C.5) can be simplified as

$$L.S. = -\int_{0}^{\frac{U_{i,i+1}-\mu_i}{\sigma_i}} \Phi(u)\,du = \int_{0}^{\frac{U_{i,i+1}-\mu_i}{\sigma_i}} \Phi(-u)\,d(-u)$$

$$= \int_{0}^{\frac{\mu_i-U_{i,i+1}}{\sigma_i}} \Phi(v)\,dv = \int_{0}^{\frac{\mu_i-U_{i,i+1}}{\sigma_i}} \Phi(u)\,du \tag{C.8}$$

Substitute this result into Equation (C.7), we get

$$\int_{0}^{\frac{\mu_i-U_{i,i+1}}{\sigma_i}} \Phi(u)\,du = \int_{0}^{\frac{U_{i,i+1}-\mu_{i+1}}{\sigma_{i+1}}} \Phi(u)\,du \tag{C.9}$$

It is obvious that Equation (C.9) exists only if

$$\frac{\mu_i - U_{i,i+1}}{\sigma_i} = \frac{U_{i,i+1} - \mu_{i+1}}{\sigma_{i+1}} \tag{C.10}$$

or

$$\mu_i \sigma_{i+1} + \mu_{i+1} \sigma_i = (\sigma_{i+1} + \sigma_i) U_{i,i+1} \tag{C.11}$$

Therefore, we have Equation (C.2).

VITA

Zhongjie Zhang was born in Shanghai, China on April 25, 1958. In 1967, he moved with his parents to Shaanxi, China. After his graduation from QingJi High School of Shaanxi in 1974, he was sent to the countryside of Shaanxi. In 1977, he was admitted by Xi'an Institute of Highways, Shaanxi and studied in civil engineering for four years. His special field at that time was "highway design and transportation engineering". He got his B.S. degree in civil engineering in 1982 and immediately entered the graduate school of Tongji University in Shanghai. He spent two and half years in the field of "soil dynamics and earthquake engineering" and obtained his M.S. in civil engineering in 1984. After his graduation he was employed by the Department of Geotechnical Engineering and the Geotechnical Engineering Research Institute in Tongji University as a teaching and research assistant (assistant professor), lecturer, and engineering consultant from 1984 to 1989. Then he was admitted by Louisiana State University in U.S.A. in 1990 and entered the PhD program of geotechnical engineering in the Department of Civil Engineering. Presently, he is a candidate for the degree of Doctor of Philosophy in Civil Engineering, specializing in Geotechnical Engineering.

DOCTORAL EXAMINATION AND DISSERTATION REPORT

Candidate: ZONGJIE ZHANG

Major Field: CIVIL ENGINEERING

Title of Dissertation: USE OF UNCERTAINTY METHODOLOGY IN IDENTIFICATION AND CLASSIFICATION OF SOILS BASED UPON CPT

Approved:

MEHMET T. TUMAY
Major Professor and Chairman

DANIEL FOGEL
Dean of the Graduate School

EXAMINING COMMITTEE:

JIANHUAN CHEN

YALCIN B. ACAR

JOSEPH N. SUHAYDA

GEORGE Z. VOYIADJIS

SUMANTA ACHARYA

Date of Examination:

OCTOBER 25, 1993